

09/889314

FILE: HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,  
JICST-EPLUS, JAPIO' ENTERED AT 10:37:11 ON 28 JAN 2003

Author(s)

L1 584 S "BURNIE J"?/AU  
L2 5349 S "MATTHEWS R"?/AU  
L3 326 S L1 AND L2  
L4 45 S (L3 OR L1 OR L2) AND PNEUMON?  
L5 21 DUP REM L4 (24 DUPLICATES REMOVED)

L5 ANSWER 1 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 1  
ACCESSION NUMBER: 2000:553692 HCAPLUS  
DOCUMENT NUMBER: 133:145931  
TITLE: Protein and DNA sequences of a novel Chlamydia  
**pneumoniae** antigen and the uses in  
diagnosis and treatment of diseases associated  
with Chlamydia infection  
INVENTOR(S): Burnie, James Peter; Matthews,  
Ruth Christine  
PATENT ASSIGNEE(S): Neutec Pharma Plc, UK  
SOURCE: PCT Int. Appl., 35 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000046359	A2	20000810	WO 2000-GB237	20000128
WO 2000046359	A3	20001207		
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1149162	A2	20011031	EP 2000-901235	20000128
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			

PRIORITY APPLN. INFO.: GB 1999-2555 A 19990205  
WO 2000-GB237 W 20000128

AB The invention provides protein and DNA sequences of a novel Chlamydia **pneumoniae** antigen. The present invention further relates to the uses of the antigens of this invention in treatment, prevention and diagnosis of infection due to Chlamydia **pneumoniae** and in particular to the prevention and treatment of atherosclerosis, including coronary atherosclerosis, caused by same.

L5 ANSWER 2 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
ACCESSION NUMBER: 1999:126671 BIOSIS  
DOCUMENT NUMBER: PREV199900126671  
TITLE: Is infection control an academic study.  
AUTHOR(S): Burnie, J. P. (1)  
CORPORATE SOURCE: (1) Dep. Med. Microbiol., 2nd Floor, Clin. Sci.  
Build., Manchester Royal Infirmary, Oxford Road,

**THIS PAGE BLANK (USPTO)**



09/889314

SOURCE: Manchester M13 9WL UK  
Journal of Hospital Infection, (Jan., 1999) Vol. 41,  
No. 1, pp. 7-10.  
ISSN: 0195-6701.

DOCUMENT TYPE: Editorial

LANGUAGE: English

AB This editorial aims to answer the question of whether infection control is an academic specialty. By considering the consequences of a lack of infection control in terms of patient morbidity and mortality and hence cost, it is easy to establish the importance of the area. Infection control embraces not only developing policies for preventing the physical spread of a microorganism but also prophylactic therapy such as vaccination and therapeutic measures such as antibiotics. Infection control not only applies to localized infection in hospital due to antibiotic resistant microbes but also to the community. Bacteria such as *Helicobacter pylori* and *Chlamydia pneumoniae* and the viruses Hepatitis B, Hepatitis C, human lymphotropic virus type 1, Epstein-Barr viruses and human papilloma virus have been implicated in diseases not previously thought to have an infectious origin. Coping with these problems is clearly an academic area.

L5 ANSWER 3 OF 21 MEDLINE DUPLICATE 2

ACCESSION NUMBER: 2000028548 MEDLINE  
DOCUMENT NUMBER: 20028548 PubMed ID: 10562808  
TITLE: A polymerase chain reaction enzyme immunoassay for diagnosing infection caused by *Aspergillus fumigatus*.  
AUTHOR: Golbang N; Burnie J P; Matthews R C  
CORPORATE SOURCE: Department of Medical Microbiology, Manchester University, Manchester Royal Infirmary, UK.  
SOURCE: JOURNAL OF CLINICAL PATHOLOGY, (1999 Jun) 52 (6) 419-23.  
Journal code: 0376601. ISSN: 0021-9746.  
PUB. COUNTRY: ENGLAND: United Kingdom  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals  
ENTRY MONTH: 199911  
ENTRY DATE: Entered STN: 20000111  
Last Updated on STN: 20000111  
Entered Medline: 19991116

AB AIM: To develop a polymerase chain reaction enzyme immunoassay (PCR-EIA) to measure levels of circulating aspergillus DNA in invasive aspergillosis caused by *Aspergillus fumigatus*. METHODS: The PCR reaction was based on primers from the 18s rRNA gene. Binding of the product to a streptavidin coated microtitration plate was mediated by a biotinylated capture probe. The product was digoxigenylated during PCR and this was the tag to which antibody was bound in the subsequent EIA. RESULTS: The optical density (OD) endpoint was < 0.1 in 10 sera from neutropenic patients with no evidence of invasive aspergillosis, and in 10 sera from nonneutropenic patients with bacterial pneumonia (group 1). The OD from five of 12 patients with allergic bronchopulmonary aspergillosis (ABPA) (group 2), three with an aspergilloma (group 3), and five with possible invasive aspergillosis (group 4) was > or = 0.1. In 63 sera from 33 cases of proven invasive aspergillosis (group 5) an OD > or = 0.1 was achieved in 48 sera from 30 patients. The maximum OD was 0.510. The level fell in survivors and gradually

**THIS PAGE BLANK (USPTO)**

09/889314

rose in fatal cases. CONCLUSIONS: This assay validated the concept of diagnosing invasive aspergillosis by measuring levels of circulating fungal DNA in serum.

L5 ANSWER 4 OF 21 MEDLINE DUPLICATE 3  
ACCESSION NUMBER: 1999135117 MEDLINE  
DOCUMENT NUMBER: 99135117 PubMed ID: 9949958  
TITLE: Is infection control an academic study?.  
AUTHOR: **Burnie J P**  
CORPORATE SOURCE: Department of Medical Microbiology, Manchester Royal Infirmary.  
SOURCE: JOURNAL OF HOSPITAL INFECTION, (1999 Jan) 41 (1) 7-10. Ref: 30  
Journal code: 8007166. ISSN: 0195-6701.  
PUB. COUNTRY: ENGLAND: United Kingdom  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
General Review; (REVIEW)  
(REVIEW, TUTORIAL)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 199904  
ENTRY DATE: Entered STN: 19990426  
Last Updated on STN: 19990426  
Entered Medline: 19990413

AB This editorial aims to answer the question of whether infection control is an academic specialty. By considering the consequences of a lack of infection control in terms of patient morbidity and mortality and hence cost, it is easy to establish the importance of the area. Infection control embraces not only developing policies for preventing the physical spread of a micro-organism but also prophylactic therapy such as vaccination and therapeutic measures such as antibiotics. Infection control not only applies to localized infection in hospital due to antibiotic resistant microbes but also to the community. Bacteria such as *Helicobacter pylori* and *Chlamydia pneumoniae* and the viruses hepatitis B, hepatitis C, human lymphotropic virus type 1, Epstein-Barr viruses and human papilloma virus have been implicated in diseases not previously thought to have an infectious origin. Coping with these problems is clearly an academic area.

L5 ANSWER 5 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.  
ACCESSION NUMBER: 97018133 EMBASE  
DOCUMENT NUMBER: 1997018133  
TITLE: Cervical cancer as an AIDS-defining illness.  
AUTHOR: Maiman M.; Fruchter R.G.; Clark M.; Arrastia C.D.; **Matthews R.**; Gates E.J.  
CORPORATE SOURCE: Dr. M. Maiman, Division of Gynecologic Oncology, Department of Obstetrics/Gynecology, SUNY-Health Science Center, 450 Clarkson Avenue, Brooklyn, NY 11203, United States  
SOURCE: Obstetrics and Gynecology, (1997) 89/1 (76-80).  
Refs: 19  
ISSN: 0029-7844 CODEN: OBGNAS  
PUBLISHER IDENT.: S 0029-7844(96)00378-X  
COUNTRY: United States  
DOCUMENT TYPE: Journal; Article  
FILE SEGMENT: 004 Microbiology  
010 Obstetrics and Gynecology

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

016 Cancer  
 017 Public Health, Social Medicine and  
 Epidemiology

LANGUAGE: English

SUMMARY LANGUAGE: English

AB Objective: To evaluate the importance of cervical cancer in the spectrum of human immunodeficiency virus (HIV) related diseases at a single high-risk institution and to compare disease characteristics in HIV-infected women with cervical cancer and those with other AIDS-related malignancies. Methods: We retrospectively reviewed data on cervical cancer and AIDS in women registered through the New York City Department of Health and institutional tumor registries from 1987 through 1995. Results: During the study period, cervical cancer was diagnosed in 28 HIV-positive women. In 26, cervical cancer was the initial AIDS-defining illness, representing 4% (26 of 725) of the subjects, and it was the sixth most common initial AIDS-defining illness in women. Cervical cancer was the most common AIDS-related malignancy among women, representing 55% of the cases, followed by lymphoma (29%) and Kaposi sarcoma (16%). In 71% of the women with cervical cancer, HIV infection was diagnosed at the time of cancer presentation by routine testing, whereas in women with other malignancies, HIV diagnosis preceded cancer diagnosis (70%) by a mean of 2.7 years. Patients with other malignancies had greater immunosuppression (mean CD4 count 153/.mu.L) than those with cervical cancer (mean CD4 count 312/.mu.L). The recurrence rate for women with cervical cancer was 88%. Although the interval from cancer diagnosis to death was similar in all three groups (9.1-12.4 months), cancer was the cause of death in 95% of HIV-infected women with cervical cancer, compared with 60% of those with other AIDS-related malignancies. Conclusion: In urban populations at increased risk for both diseases, cervical cancer is an important AIDS-defining illness and may be the most common AIDS-related malignancy in women.

L5 ANSWER 6 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 1996:109315 BIOSIS

DOCUMENT NUMBER: PREV199698681450

TITLE: Immunoblotting in the diagnosis of culture negative endocarditis caused by streptococci and enterococci.

AUTHOR(S): Burnie, J. P. (1); Clark, I.

CORPORATE SOURCE: (1) Dep. Medical Microbiol., 2nd Floor, Clinical Sci. Build., Manchester Royal Infirmary, Oxford Road, Manchester M13 9WL UK

SOURCE: Journal of Clinical Pathology (London), (1995) Vol. 48, No. 12, pp. 1130-1136.  
 ISSN: 0021-9746.

DOCUMENT TYPE: Article

LANGUAGE: English

AB Aim: To improve the diagnosis of culture negative endocarditis by diagnosing cases due to streptococci and enterococci. Method: Serum samples were immunoblotted against extracts of the commonest streptococci and enterococci. They were selected from patients with a cardiac murmur, persistent pyrexia and at least three negative blood cultures. The presence of patterns of endocarditis species specific antigenic bands was measured and correlated with clinical outcome. Results: Negative serology was found in 28 patients where the diagnosis of endocarditis was rejected or, if proved, staphylococcal, yeast, Gram negative, systemic lupus erythematosus,

**THIS PAGE BLANK (USPTO)**

due to Q fever or Chlamydia psittaci or nonbacterial thrombotic. Positive serology was found in 27 of the 34 patients where the response to antibiotics suggested streptococcal or enterococcal infection. In 22 of these there was objective evidence of endocarditis. Positive serology was also found in three of four further patients with vegetations at necropsy. Conclusion: The identification of patterns of antibody response on immunoblotting can be used to make a specific diagnosis of streptococcal or enterococcal endocarditis in the absence of positive blood cultures.

L5 ANSWER 7 OF 21 SCISEARCH COPYRIGHT 2003 ISI (R)  
 ACCESSION NUMBER: 94:539953 SCISEARCH  
 THE GENUINE ARTICLE: PE730  
 TITLE: THE LEUCINE-RESPONSIVE REGULATORY PROTEIN, A GLOBAL  
 REGULATOR OF METABOLISM IN ESCHERICHIA-COLI  
 AUTHOR: CALVO J M (Reprint); MATTHEWS R G  
 CORPORATE SOURCE: CORNELL UNIV, BIOCHEM MOLEC & CELL BIOL SECT,  
 ITHACA, NY, 14853 (Reprint); UNIV MICHIGAN, DIV  
 BIOPHYS RES, ANN ARBOR, MI, 48109; UNIV MICHIGAN,  
 DEPT BIOL CHEM, ANN ARBOR, MI, 48109  
 COUNTRY OF AUTHOR: USA  
 SOURCE: MICROBIOLOGICAL REVIEWS, (SEP 1994) Vol. 58, No. 3,  
 pp. 466-490.  
 ISSN: 0146-0749.  
 DOCUMENT TYPE: General Review; Journal  
 FILE SEGMENT: LIFE  
 LANGUAGE: ENGLISH  
 REFERENCE COUNT: 187

\*ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS\*

AB Natural genetic transformation is the active uptake of free DNA by bacterial cells and the heritable incorporation of its genetic information. Since the famous discovery of transformation in *Streptococcus pneumoniae* by Griffith in 1928 and the demonstration of DNA as the transforming principle by Avery and coworkers in 1944, cellular processes involved in transformation have been studied extensively by in vitro experimentation with a few transformable species. Only more recently has it been considered that transformation may be a powerful mechanism of horizontal gene transfer in natural bacterial populations. In this review the current understanding of the biology of transformation is summarized to provide the platform on which aspects of bacterial transformation in water soil, and sediments and the habitat of pathogens are discussed. Direct and indirect evidence for gene transfer routes by transformation within species and between different species will be presented, along with data suggesting that plasmids as well as chromosomal DNA are subject to genetic exchange via transformation. Experiments exploring the prerequisites for transformation in the environment, including the production and persistence of free DNA and factors important for the uptake of DNA by cells, will be compiled, as well as possible natural barriers to transformation. The efficiency of gene transfer by transformation in bacterial habitats is possibly genetically adjusted to submaximal levels. The fact that natural transformation has been detected among bacteria from all trophic and taxonomic groups including archaeobacteria suggests that transformability evolved early in phylogeny. Probable functions of DNA uptake other than gene acquisition will be discussed. The body of information presently available suggests that transformation has a great impact on bacterial population dynamics

**THIS PAGE BLANK (USPTO)**

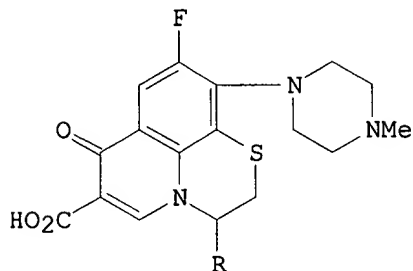


09/889314

as well as on bacterial evolution and speciation.

L5 ANSWER 8 OF 21 SCISEARCH COPYRIGHT 2003 ISI (R)  
ACCESSION NUMBER: 92:717086 SCISEARCH  
THE GENUINE ARTICLE: KA952  
TITLE: PROGRESSIVE CHRONIC PULMONARY ASPERGILLOSIS - A  
DIAGNOSTIC AND THERAPEUTIC CHALLENGE  
AUTHOR: ELLIS M E (Reprint); DOSSING M; ALHOKAIL A; QADRI S  
H; HAINAU B; **BURNIE J**  
CORPORATE SOURCE: KING FAISAL SPECIALIST HOSP & RES CTR, DEPT MED, MBC  
46, POB 3354, RIYADH 11211, SAUDI ARABIA (Reprint);  
KING FAISAL SPECIALIST HOSP & RES CTR, DEPT PATHOL &  
LAB MED, RIYADH 11211, SAUDI ARABIA; UNIV  
MANCHESTER, SCH MED, DEPT MICROBIOL, MANCHESTER M13  
9PL, LANCs, ENGLAND  
COUNTRY OF AUTHOR: SAUDI ARABIA; ENGLAND  
SOURCE: JOURNAL OF THE ROYAL SOCIETY OF MEDICINE, (DEC 1992)  
Vol. 85, No. 12, pp. 763-764.  
ISSN: 0141-0768.  
DOCUMENT TYPE: Article; Journal  
FILE SEGMENT: LIFE; CLIN  
LANGUAGE: ENGLISH  
REFERENCE COUNT: 13

L5 ANSWER 9 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 4  
ACCESSION NUMBER: 1992:507995 HCAPLUS  
DOCUMENT NUMBER: 117:107995  
TITLE: The in vitro activity of two new quinolones:  
rufloxacin and MF 961  
AUTHOR(S): Wise, R.; Andrews, J. M.; **Matthews, R.**  
; Wolstenholme, M.  
CORPORATE SOURCE: Dep. Med. Microbiol., Dudley Road Hosp.,  
Birmingham, B18 7QH, UK  
SOURCE: Journal of Antimicrobial Chemotherapy (1992),  
29(6), 649-60  
CODEN: JACHDX; ISSN: 0305-7453  
DOCUMENT TYPE: Journal  
LANGUAGE: English  
GI



I, R=H

II, R=CH<sub>2</sub>F

AB The in vitro activity of two new quinolone antimicrobials,  
rufloxacin (I) and MF 961 (II), together with the desmethylated  
metabolite of rufloxacin (MF 922) were compared with other orally  
administered agents against 622 bacterial strains. Against

**THIS PAGE BLANK (USPTO)**

09/889314

Enterobacteriaceae and *Pseudomonas aeruginosa*, I was generally active (MIC90 1-8 mg/L) with the exception of *Klebsiella* and *Serratia* spp. (MIC90 32 mg/L) and *Enterobacter* spp. (MIC90 64 mg/L). The respiratory pathogens *Haemophilus influenzae* and *Moraxella catarrhalis* were susceptible to I (MIC90 0.5 and 1 mg/L, resp.), but *Streptococcus pneumoniae* was less susceptible (MIC90 32 mg/L). *Staphylococcus aureus* was susceptible to I (MIC90 2 mg/L). The I metabolite, MF 922, was generally as active as its parent. II was usually 2-fold more active than I. All 3 compds. were 4-16 times less active than norfloxacin, but I was as active or somewhat more active than norfloxacin against *Staphylococcus* spp. Any strains showing decreased susceptibility to other quinolones exhibited cross resistance to these new agents. The MBC of I and MF 922 was within one diln. of the MIC, and human serum had little effect upon the activity of both agents. The protein binding of I and MF 922 at 1 and 10 mg/L was 55 and 63.8% and 30.3% and 32.6%, resp. The activity of rifloxacin against four strains of *Chlamydia trachomatis* and one strain of *Chlamydia pneumoniae* was detd. The MIC for *C. trachomatis* was 4-8 mg/L and 4 mg/L for *C. pneumoniae*.

L5 ANSWER 10 OF 21 MEDLINE  
ACCESSION NUMBER: 92115389 MEDLINE  
DOCUMENT NUMBER: 92115389 PubMed ID: 1766717  
TITLE: *Chlamydia pneumoniae* (TWAR) in neonates.  
AUTHOR: Matthews R S; Mohite A; Addy D P; Wise R  
SOURCE: PEDIATRIC INFECTIOUS DISEASE JOURNAL, (1991 Dec) 10  
(12) 956-7.  
Journal code: 8701858. ISSN: 0891-3668.  
PUB. COUNTRY: United States  
DOCUMENT TYPE: Letter  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 199202  
ENTRY DATE: Entered STN: 19920308  
Last Updated on STN: 19920308  
Entered Medline: 19920214

L5 ANSWER 11 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
DUPLICATE 5  
ACCESSION NUMBER: 1991:226713 BIOSIS  
DOCUMENT NUMBER: BA91:118173  
TITLE: ISOLATION CLONING MAPPING AND NUCLEOTIDE SEQUENCING  
OF THE GENE ENCODING FLAVODOXIN IN *ESCHERICHIA-COLI*.  
AUTHOR(S): OSBORNE C; CHEN L-M; MATTHEWS R G  
CORPORATE SOURCE: DEP. BIOL. CHEM., UNIV. MICHIGAN, ANN ARBOR, MICH.  
48109.  
SOURCE: J BACTERIOL, (1991) 173 (5), 1729-1737.  
CODEN: JOBAAY. ISSN: 0021-9193.  
FILE SEGMENT: BA; OLD  
LANGUAGE: English

AB The flavodoxins constitute a highly conserved family of small, acidic electron transfer proteins with flavin mononucleotide prosthetic groups. They are found in prokaryotes and in red and green algae, where they provide electron at low potentials for the reduction of nitrogen by nitrogenase, for the light-dependent reduction of NADP+ in photosynthesis, and for the reduction of sulfite. Proteins with the physical characteristics of flavodoxins

**THIS PAGE BLANK (USPTO)**

have been implicated in the reductive activation of pyruvate formate-lyase and cobalamin-dependent methionine synthase in *Escherichia coli*. We have purified flavodoxin to homogeneity from *E. coli*, determined its N-terminal amino acid sequence, and used this sequence to construct a 64-fold degenerate oligonucleotide probe for the flavodoxin gene. Because the phenotype of a flavodoxin mutant is not known, we used this degenerate probe to screen the phages of Kohara library and identified two phages, with inserts mapping at .apprx. 16 min, that hybridized to the probe. The flavodoxin gene, designated *fldA*, was subcloned from the DNA in the overlap region of these two clones. The deduced amino acid sequence, determined by nucleotide sequencing of the flavodoxin gene, shows strong homology with flavodoxins from nitrogen-fixing bacteria and cyanobacteria. The *fldA* gene maps at 15.9 min on the *E. coli* chromosome and is transcribed in a counterclockwise direction.

L5 ANSWER 12 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.DUPLICATE  
6

ACCESSION NUMBER: 92041225 EMBASE  
DOCUMENT NUMBER: 1992041225  
TITLE: Chlamydia **pneumoniae** (TWAR) in neonates [1].  
AUTHOR: **Matthews R.S.**; Mohite A.; Addy D.P.; Wise R.  
CORPORATE SOURCE: Dudley Road Hospital, Birmingham, United Kingdom  
SOURCE: Pediatric Infectious Disease Journal, (1991) 10/12 (956-957).  
ISSN: 0891-3668 CODEN: PIDJEV  
COUNTRY: United States  
DOCUMENT TYPE: Journal; Letter  
FILE SEGMENT: 004 Microbiology  
007 Pediatrics and Pediatric Surgery  
030 Pharmacology  
037 Drug Literature Index  
LANGUAGE: English

L5 ANSWER 13 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 7  
ACCESSION NUMBER: 1991:651968 HCAPLUS  
DOCUMENT NUMBER: 115:251968  
TITLE: In-vitro susceptibility of Chlamydia **pneumoniae** (TWAR) to seven antibiotics  
AUTHOR(S): Cooper, M. A.; Baldwin, D.; **Matthews, R. S.**; Andrews, J. M.; Wise, R.  
CORPORATE SOURCE: Dep. Microbiol., Dudley Road Hosp., Birmingham, UK  
SOURCE: Journal of Antimicrobial Chemotherapy (1991), 28(3), 407-13  
CODEN: JACHDX; ISSN: 0305-7453  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB A modification of an immunofluorescence method previously used to study the in vitro antimicrobial susceptibilities of *C. trachomatis* was used to investigate the activity of seven antimicrobials against a strain of *C. pneumoniae*. Adding antimicrobial to preinfected cells gave higher MICs (min. inhibitory concns.) and MLCs (min. lethal concns.) than when cells were infected in the presence of the antimicrobials, and this difference in methodol. could account for the discrepancy between these results and those of

**THIS PAGE BLANK (USPTO)**

09/889314

others. Of the antimicrobials studied, clarithromycin and its 14-hydroxy metabolite were the most active agents; sparfloxacin was more active than ciprofloxacin, but no more active than more conventional antichlamydial agents.

L5 ANSWER 14 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.  
ACCESSION NUMBER: 1991:535104 BIOSIS  
DOCUMENT NUMBER: BR41:124839  
TITLE: THE ACTIVITY OF SPARFLOXACIN AND OTHER AGENTS AGAINST CHLAMYDIA-**PNEUMONIAE** THE EFFECT OF PRE AND POST ANTIBIOTIC EXPOSURE.  
AUTHOR(S): WISE R; ANDREWS J M; COOPER M A; **MATTHEWS R**  
CORPORATE SOURCE: DUDLEY ROAD HOSP., BIRMINGHAM B18 7QH, UK.  
SOURCE: THIRTY-FIRST INTERSCIENCE CONFERENCE ON ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, CHICAGO, ILLINOIS, USA, SEPTEMBER 29-OCTOBER 2, 1991. PROGRAM ABSTR, (1991) 31 (0), 212.  
CODEN: POCHES.  
DOCUMENT TYPE: Conference  
FILE SEGMENT: BR; OLD  
LANGUAGE: English

L5 ANSWER 15 OF 21 MEDLINE DUPLICATE 8  
ACCESSION NUMBER: 91324522 MEDLINE  
DOCUMENT NUMBER: 91324522 PubMed ID: 1907617  
TITLE: Immunoblotting and culture positive endocarditis.  
AUTHOR: Clark I; **Burnie J P**  
CORPORATE SOURCE: Department of Medical Microbiology, University of Manchester Medical School.  
SOURCE: JOURNAL OF CLINICAL PATHOLOGY, (1991 Feb) 44 (2) 152-6.  
Journal code: 0376601. ISSN: 0021-9746.  
PUB. COUNTRY: ENGLAND: United Kingdom  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals  
ENTRY MONTH: 199109  
ENTRY DATE: Entered STN: 19910929  
Last Updated on STN: 19910929  
Entered Medline: 19910909

AB Serum samples from patients with endocarditis due to Streptococcus mutans, Streptococcus **pneumoniae**, Streptococcus agalactiae, Streptococcus lactis and a "nutritionally" variant streptococcus were immunoblotted against antigenic extracts from all five species. In S mutans endocarditis there was an endocarditis specific pattern of IgM against bands of 220, 200, and 190 kilodaltons. In S **pneumoniae** IgM against antibody of a molecular weight greater than 150 kilodaltons was specific to endocarditis. In S agalactiae IgM against bands at 82, 71, and 66-67 kilodaltons was endocarditis specific. In S lactis endocarditis specific IgM was present against antigenic bands at 105, 66, 61 and 58 kilodaltons. With the "nutritionally" variant streptococcus it was impossible to distinguish between cases of endocarditis and controls.

L5 ANSWER 16 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 9  
ACCESSION NUMBER: 1991:58770 HCAPLUS  
DOCUMENT NUMBER: 114:58770

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**



09/889314

TITLE: In vitro activity of sparfloxacin, a new  
quinolone antimicrobial agent  
AUTHOR(S): Cooper, M. A.; Andrews, J. M.; Ashby, J. P.;  
**Matthews, R. S.**; Wise, R.  
CORPORATE SOURCE: Dep. Microbiol., Dudley Road Hosp., Birmingham,  
UK  
SOURCE: Journal of Antimicrobial Chemotherapy (1990),  
26(5), 667-76  
CODEN: JACHDX; ISSN: 0305-7453  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB The in vitro activity of sparfloxacin (AT-4140), a new difluorinated quinolone, was compared with those of ciprofloxacin, temafloxacin, and selected members of other groups of antimicrobial agents, against 651 recent distinct clin. isolates and strains with known mechanisms of resistance. Three strains of *Chlamydia trachomatis* were also studied. The MICs for 90% of the Enterobacteriaceae were 0.06-1 mg/L; for *Pseudomonas aeruginosa* the MIC<sub>90</sub> was 2 mg/L. Sparfloxacin was 16-fold more active against *Acinetobacter* spp. than ciprofloxacin. For *Staphylococcus* spp., *Streptococcus* spp., and *Enterococcus faecalis*, the MIC<sub>90</sub> was 0.25-1 mg/L; sparfloxacin was 4-fold more active against *Streptococcus pneumoniae* than ciprofloxacin. Ninety percent of strains of *Haemophilus influenzae*, *Branhamella catarrhalis*, and *Neisseria* spp. were inhibited by <0.03 mg/L; for *Bacteroides fragilis* the MIC<sub>90</sub> was 1 mg/L. The 3 strains of *C. trachomatis* were susceptible to 0.06-0.12 mg/L sparfloxacin, which was 16-fold more active than ciprofloxacin. There was cross resistance among the quinolones, but not between the quinolones and other groups of antimicrobials. The protein binding of sparfloxacin was 40% and serum had little effect on its activity.

L5 ANSWER 17 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 10

ACCESSION NUMBER: 1988:435176 HCAPLUS  
DOCUMENT NUMBER: 109:35176  
TITLE: In vitro activity of lomefloxacin, a new  
quinolone antimicrobial agent, in comparison  
with those for other agents  
AUTHOR(S): Wise, R.; Andrews, J. M.; Ashby, J. P.;  
**Matthews, R. S.**  
CORPORATE SOURCE: Dep. Med. Microbiol., Dudley Road Hosp.,  
Birmingham, B18 7QH, UK  
SOURCE: Antimicrobial Agents and Chemotherapy (1988),  
32(5), 617-22  
CODEN: AMACQ; ISSN: 0066-4804  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB The in vitro activity of lomefloxacin (I), a new difluorinated quinolone, was compared with those of ofloxacin, ciprofloxacin, fleroxacin, amoxicillin, cefuroxime, and trimethoprim against 585 recent clin. isolates and other strains with known mechanisms of resistance. The MICs of I against 90% of the members of the family Enterobacteriaceae, *Pseudomonas aeruginosa*, and staphylococci were between 0.25 and 4 .mu.g/mL. Ninety percent of *Neisseria* sp. and *Haemophilus influenzae* were susceptible to .ltoreq.0.06 .mu.g/mL, and streptococci (including *S. pyogenes*, *S. pneumoniae*, and enterococci) and *Bacteroides feragilis* were susceptible to 8 .mu.g/mL. I was comparable in activity to fleroxacin and ofloxacin, but it was less active than ciprofloxacin. There was

**THIS PAGE BLANK (USPTO)**

09/889314

cross-resistance between the quinolone group of antimicrobial agents. The protein binding of I was 15.4%, and serum had little effect on the activity of the compd. However, urine at pH 5.0 decreased the activity by 2-8-fold compared with that at pH 7.0.

L5 ANSWER 18 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 88049469 EMBASE

DOCUMENT NUMBER: 1988049469

TITLE: Chlamydial rhinitis neonatorum.

AUTHOR: Kent S.E.; **Matthews R.S.**

CORPORATE SOURCE: Birmingham Children's Hospital, Ladywood, Birmingham B16 8ET, United Kingdom

SOURCE: Journal of Laryngology and Otology, (1987) 101/11 (1193-1197).

ISSN: 0022-2151 CODEN: JLOTAX

COUNTRY: United Kingdom

DOCUMENT TYPE: Journal

FILE SEGMENT: 004 Microbiology  
007 Pediatrics and Pediatric Surgery  
010 Obstetrics and Gynecology  
011 Otorhinolaryngology

LANGUAGE: English

SUMMARY LANGUAGE: English

AB This paper reports the case of a five-week-old child who developed severe rhinitis with epistaxes, due to infection with chlamydia trachomatis. The infection was acquired from the mother's genital tract during birth but the true diagnosis was not suspected for some time. Chlamydia trachomatis is well recognized as one of the causes of neonatal conjunctivitis and **pneumonia**. However, its presence in the upper respiratory tract is thought usually to be asymptomatic (Oriol and Ridgway, 1982). Little attention has been focused on its role in neonatal rhinitis. This paper draws to the otolaryngologist's attention the entity of chlamydial rhinitis neonatorum and discusses the epidemiology, diagnosis and management of the condition.

L5 ANSWER 19 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 85172703 EMBASE

DOCUMENT NUMBER: 1985172703

TITLE: CM-40874.

AUTHOR: **Burnie J.; Matthews R.**

CORPORATE SOURCE: United Kingdom

SOURCE: Drugs of the Future, (1985) 10/3 (193-195).

CODEN: DRFUD4

COUNTRY: Spain

DOCUMENT TYPE: Journal

FILE SEGMENT: 037 Drug Literature Index  
052 Toxicology

LANGUAGE: English

L5 ANSWER 20 OF 21 MEDLINE

ACCESSION NUMBER: 78198713 MEDLINE

DOCUMENT NUMBER: 78198713 PubMed ID: 78425

TITLE: Multiply resistant pneumococcus.

AUTHOR: Meers P D; **Matthews R B**

SOURCE: LANCET, (1978 Jul 22) 2 (8082) 219.

Journal code: 2985213R. ISSN: 0140-6736.

PUB. COUNTRY: ENGLAND: United Kingdom

**THIS PAGE BLANK (USPTO)**

09/889314

DOCUMENT TYPE: Letter  
LANGUAGE: English  
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals  
ENTRY MONTH: 197809  
ENTRY DATE: Entered STN: 19900314  
Last Updated on STN: 19950206  
Entered Medline: 19780901

L5 ANSWER 21 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.  
ACCESSION NUMBER: 78377563 EMBASE  
DOCUMENT NUMBER: 1978377563  
TITLE: Multiply resistant pneumococcus.  
AUTHOR: Meers P.D.; **Matthews R.B.**  
CORPORATE SOURCE: Publ. Hlth Lab., Gen. Hosp., Plymouth PL4 8NN, United Kingdom  
SOURCE: Lancet, (1978) 2/8081 (155-156).  
CODEN: LANCAO  
COUNTRY: United Kingdom  
DOCUMENT TYPE: Journal  
FILE SEGMENT: 037 Drug Literature Index  
004 Microbiology  
LANGUAGE: English

AB A pneumococcus resistant to chloramphenicol, tetracycline and penicillin was isolated from a holidaymaker from Spain, who was admitted to hospital in Plymouth with an acute skin condition. She had a productive cough, and from mucopurulent sputum a pneumococcus (type 23, Danish nomenclature) was isolated. On routine testing this was found to be resistant to chloramphenicol and tetracycline, and the zone of inhibition round the penicillin disc (1.5 units) was reduced from the expected .gtoreq.10 mm to 7 mm, measured radially from the edge of the disc. By the agar dilution method, using a minimal-growth endpoint, the minimum inhibitory concentration (M.I.C.) of penicillin was 0.25 .mu.g/ml compared with an M.I.C. of 0.03 .mu.g/ml for a sensitive strain. The M.I.C.s of tetracycline and chloramphenicol were both 20 .mu.g/ml. The inoculum used in these tests was about 106 organisms.

FILE 'HOME' ENTERED AT 10:38:23 ON 28 JAN 2003

19 18 17 16 15

**THIS PAGE BLANK (USPTO)**

977 6 1.2 515 1 US-08-063-552-4 Sequence 4, Appli  
 978 6 1.2 515 5 PCT-US93-05704-4 Patent No. RE34,60  
 979 6 1.2 522 6 RE34606-6 Sequence 2, Appli  
 980 6 1.2 525 1 US-08-348-891A-2 Sequence 2, Appli  
 981 6 1.2 525 2 US-08-905-817-2 Sequence 8, Appli  
 982 6 1.2 525 2 US-09-351-224E-8 Sequence 40, Appli  
 983 6 1.2 526 2 US-08-853-659A-40 Sequence 2, Appli  
 984 6 1.2 529 1 US-08-548-509-2 Sequence 88, Appli  
 985 6 1.2 532 4 US-09-071-035-88 Sequence 49, Appli  
 986 6 1.2 535 1 US-07-872-644-49 Sequence 49, Appli  
 987 6 1.2 535 1 US-08-297-510-49 Sequence 49, Appli  
 988 6 1.2 535 1 US-08-479-532-49 Sequence 49, Appli  
 989 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli  
 990 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli  
 991 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli  
 992 6 1.2 535 3 US-08-813-574-2 Sequence 2, Appli  
 993 6 1.2 535 3 US-09-139-491-49 Sequence 49, Appli  
 994 6 1.2 535 5 PCT-US92-03222-49 Sequence 49, Appli  
 995 6 1.2 540 4 US-09-011-074-4 Sequence 17, Appli  
 996 6 1.2 542 4 US-09-380-420C-17 Sequence 3, Appli  
 997 6 1.2 545 4 US-09-351-224E-3 Sequence 2, Appli  
 998 6 1.2 550 1 US-08-484-493-2 Sequence 2, Appli  
 999 6 1.2 550 2 US-08-484-494-2 Sequence 2, Appli  
 1000 6 1.2 550 2 US-08-345-212-2

## ALIGNMENTS

102 (e)

RESULT 1  
 US-08-809-326A-1  
 Sequence 1, Application US/08809326A

Patent No. 6165478  
 GENERAL INFORMATION:  
 APPLICANT: Iutsu, Hiroshi  
 APPLICANT: Obata, Kazuhiko  
 APPLICANT: Matsumoto, Akira  
 TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
 FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMANTS  
 TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
 TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
 TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
 TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR  
 TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 TITLE OF INVENTION: PNEUMONIAE GENE  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036-2711

COMPUTER READABLE FORM:  
 MEDIAN TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/809,326A  
 FILING DATE: 19-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 224711/94  
 FILING DATE: 20-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106006/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106008/95  
 FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106009/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106010/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106011/95  
 FILING DATE: 28-APR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Miller, Charles E.  
 REGISTRATION NUMBER: 24,576  
 REFERENCE/DOCKET NUMBER: 7426-043-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 488 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-809-326A-1

Query Match 98.2%; Score 487; DB 4; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSVLTSTPGVPOODKLSGNETKOIOFROGKNTMESDATTAG 63  
 DB 1 MSISSSGPDNOKNIMSVLTSTPGVPOODKLSGNETKOIOFROGKNTMESDATTAG 60  
 QY 64 ASGKDKTSTTETETAPPOGVAAGKESSESOKAGADTVSGAAATTAATKIAMOTSI 123  
 DB 61 ASGKDKTSTTETETAPPOGVAAGKESSESOKAGADTVSGAAATTAATKIAMOTSI 120  
 QY 124 EASMSMESTLESLOSLSAOKKEVEAYVVAALSGKSSGSALETPELPKGVPRSEVI 183  
 DB 121 EASMSMESTLESLOSLSAOKKEVEAYVVAALSGKSSGSALETPELPKGVPRSEVI 180  
 QY 184 EIGLALRAIOTLGATFSALSNTASTOQADQTKLGLKQAIKIDEREYQEMKAAE 243  
 DB 181 EIGLALRAIOTLGATFSALSNTASTOQADQTKLGLKQAIKIDEREYQEMKAAE 240  
 QY 244 OKSKDEGTMDTVNTVMIAVSAITVISIAAIFTCGAGLAGLAAGAAVGAAGAAGA 303  
 DB 241 OKSKDEGTMDTVNTVMIAVSAITVISIAAIFTCGAGLAGLAAGAAVGAAGAAGA 300  
 QY 304 AATTVAAQITVQAVVQAVKAVITAVQAATTAIKAAVSGIKAFITVYKAIAKATSK 363  
 DB 301 AATTVAAQITVQAVVQAVKAVITAVQAATTAIKAAVSGIKAFITVYKAIAKATSK 360  
 QY 364 GISKVFAGKTOMIAKNFPLSKVISLSKVTYVGVVVAAPALGKIGIMQOLSEMOON 423  
 DB 361 GISKVFAGKTOMIAKNFPLSKVISLSKVTYVGVVVAAPALGKIGIMQOLSEMOON 420  
 QY 424 VAQFOKEVGKLOAAADMTSMFTQFMOQASKIASKOTGESNEMTORATKLGAIILKAYAI 483  
 DB 421 VAQFOKEVGKLOAAADMTSMFTQFMOQASKIASKOTGESNEMTORATKLGAIILKAYAI 480  
 QY 484 SGATIGA 490  
 DB 481 SGATIGA 487

RESULT 2  
 US-08-809-326A-15  
 Sequence 15, Application US/08809326A  
 Patent No. 6165478  
 GENERAL INFORMATION:  
 APPLICANT: Iutsu, Hiroshi  
 APPLICANT: Obata, Kazuhiko

Tue Jan 28

**THIS PAGE BLANK (UBPTD)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: January 27, 2003, 16:35:23 : Search time 15 Seconds  
(without alignments)  
972.918 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496

Sequence: 1 DNNMSSSSSGPDNOKNIMS.....LKAYAIISGAIAGAKHTNMF 496

## Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

## Post-processing: Listing first 1000 summaries

## Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	98.2	488	US-08-809-326A-1	Sequence 1, Appl
2	487	98.2	649	US-08-809-326A-15	Sequence 15, Appl
3	259	52.2	259	US-08-809-326A-5	Sequence 5, Appl
4	259	52.2	271	US-08-809-326A-2	Sequence 2, Appl
5	259	52.2	432	US-08-809-326A-16	Sequence 16, Appl
6	8	1.6	27	US-09-135-994-11	Sequence 11, Appl
7	8	1.6	28	US-08-425-069-50	Sequence 50, Appl
8	8	1.6	28	US-08-317-844B-50	Sequence 50, Appl
9	8	1.6	30	US-08-425-069-61	Sequence 61, Appl
10	8	1.6	30	US-08-317-844B-61	Sequence 61, Appl
11	8	1.6	33	US-08-425-069-6	Sequence 6, Appl
12	8	1.6	33	US-08-317-844B-6	Sequence 6, Appl
13	8	1.6	111	US-08-185-414E-2	Sequence 2, Appl
14	8	1.6	129	US-09-135-994-12	Sequence 12, Appl
15	8	1.6	164	US-08-911-319A-1	Sequence 1, Appl
16	8	1.6	164	US-09-352-619-1	Sequence 1, Appl
17	8	1.6	205	US-08-775-009-37	Sequence 37, Appl
18	8	1.6	331	US-08-364-081-3	Sequence 3, Appl
19	8	1.6	331	US-08-630-552-3	Sequence 3, Appl
20	8	1.6	331	US-08-16558-3	Sequence 3, Appl
21	8	1.6	625	US-09-347-801-18	Sequence 18, Appl
22	8	1.6	651	US-08-556-978B-19	Sequence 19, Appl
23	8	1.6	651	US-09-247-806-1	Sequence 1, Appl
24	8	1.6	718	US-08-425-069-2	Sequence 2, Appl
25	8	1.6	718	US-08-317-844B-2	Sequence 2, Appl
26	8	1.6	747	US-09-034-177-3	Sequence 3, Appl
27	8	1.6	1400	US-08-630-915A-37	Sequence 37, Appl

28	8	1.6	2088	US-09-548-372D-13	Sequence 13, Appl
29	8	1.6	2088	US-09-548-367D-13	Sequence 13, Appl
30	8	1.6	2211	US-09-738-884-1	Sequence 1, Appl
31	8	1.6	2308	US-09-627-650B-7	Sequence 7, Appl
32	8	1.6	2508	US-09-436-063C-7	Sequence 7, Appl
33	8	1.6	2544	US-09-627-650B-3	Sequence 3, Appl
34	8	1.6	2544	US-09-436-063C-3	Sequence 3, Appl
35	8	1.6	2601	US-09-627-650B-9	Sequence 9, Appl
36	8	1.6	2601	US-09-436-063C-9	Sequence 9, Appl
37	8	1.6	3788	US-09-336-447A-76	Sequence 76, Appl
38	8	1.4	11	5178861-16	Patent No. 5178861
39	7	1.4	21	US-08-425-069-5	Sequence 5, Appl
40	7	1.4	21	US-08-317-844B-5	Sequence 5, Appl
41	7	1.4	23	US-07-987-286-3	Sequence 3, Appl
42	7	1.4	23	US-08-614-626-3	Sequence 3, Appl
43	7	1.4	25	US-07-987-286-11	Sequence 11, Appl
44	7	1.4	25	US-07-987-286-24	Sequence 24, Appl
45	7	1.4	25	US-08-425-069-38	Sequence 38, Appl
46	7	1.4	25	US-08-614-626-11	Sequence 11, Appl
47	7	1.4	25	US-08-614-626-24	Sequence 24, Appl
48	7	1.4	25	US-08-317-844B-38	Sequence 38, Appl
49	7	1.4	27	US-08-425-069-47	Sequence 47, Appl
50	7	1.4	27	US-08-425-069-51	Sequence 51, Appl
51	7	1.4	27	US-08-425-069-55	Sequence 55, Appl
52	7	1.4	27	US-08-317-844B-47	Sequence 47, Appl
53	7	1.4	27	US-08-317-844B-51	Sequence 51, Appl
54	7	1.4	27	US-08-317-844B-55	Sequence 55, Appl
55	7	1.4	27	US-08-556-978B-29	Sequence 29, Appl
56	7	1.4	27	US-08-556-978B-92	Sequence 92, Appl
57	7	1.4	28	US-08-425-069-40	Sequence 40, Appl
58	7	1.4	28	US-08-425-069-43	Sequence 43, Appl
59	7	1.4	28	US-08-317-844B-40	Sequence 40, Appl
60	7	1.4	28	US-08-317-844B-43	Sequence 43, Appl
61	7	1.4	30	US-08-425-069-41	Sequence 41, Appl
62	7	1.4	30	US-08-425-069-48	Sequence 48, Appl
63	7	1.4	30	US-08-425-069-53	Sequence 53, Appl
64	7	1.4	30	US-08-425-069-58	Sequence 58, Appl
65	7	1.4	30	US-08-425-069-60	Sequence 60, Appl
66	7	1.4	30	US-08-317-844B-41	Sequence 41, Appl
67	7	1.4	30	US-08-317-844B-48	Sequence 48, Appl
68	7	1.4	30	US-08-317-844B-53	Sequence 53, Appl
69	7	1.4	30	US-08-317-844B-58	Sequence 58, Appl
70	7	1.4	30	US-08-317-844B-60	Sequence 60, Appl
71	7	1.4	30	US-08-556-978B-32	Sequence 32, Appl
72	7	1.4	30	US-08-556-978B-75	Sequence 75, Appl
73	7	1.4	30	US-08-556-978B-95	Sequence 95, Appl
74	7	1.4	31	US-08-425-069-46	Sequence 46, Appl
75	7	1.4	31	US-08-317-844B-46	Sequence 46, Appl
76	7	1.4	31	US-08-556-978B-26	Sequence 26, Appl
77	7	1.4	31	US-08-556-978B-69	Sequence 69, Appl
78	7	1.4	31	US-08-556-978B-89	Sequence 89, Appl
79	7	1.4	32	US-08-425-069-44	Sequence 44, Appl
80	7	1.4	32	US-08-317-844B-44	Sequence 44, Appl
81	7	1.4	33	US-08-425-069-7	Sequence 7, Appl
82	7	1.4	33	US-08-317-844B-7	Sequence 7, Appl
83	7	1.4	34	US-08-425-069-42	Sequence 42, Appl
84	7	1.4	34	US-08-425-069-52	Sequence 52, Appl
85	7	1.4	34	US-08-425-069-54	Sequence 54, Appl
86	7	1.4	34	US-08-317-844B-42	Sequence 42, Appl
87	7	1.4	34	US-08-317-844B-52	Sequence 52, Appl
88	7	1.4	34	US-08-317-844B-54	Sequence 54, Appl
89	7	1.4	58	US-08-279-058B-23	Sequence 23, Appl
90	7	1.4	58	US-08-828-323-23	Sequence 23, Appl
91	7	1.4	76	5496550-6	Patent No. 5496550
92	7	1.4	85	US-08-244-116B-13	Sequence 13, Appl
93	7	1.4	101	US-08-556-978B-20	Sequence 20, Appl
94	7	1.4	101	US-08-556-978B-22	Sequence 22, Appl
95	7	1.4	101	US-08-556-978B-62	Sequence 62, Appl
96	7	1.4	101	US-09-247-806-3	Sequence 3, Appl
97	7	1.4	101	US-09-247-806-5	Sequence 5, Appl
98	7	1.4	101	US-09-247-806-7	Sequence 7, Appl
99	7	1.4	102	US-08-273-146-65	Sequence 65, Appl
100	7	1.4	102	US-08-974-022-53	Sequence 53, Appl

101	7	1.4	102	4	US-08-795-445A-53	Sequence 53, Appl	174	7	1.4	445	3	US-09-165-543-5	Sequence 5, Appl
102	7	1.4	102	4	US-08-795-447A-53	Sequence 53, Appl	175	7	1.4	445	3	US-09-167-354-7	Sequence 7, Appl
103	7	1.4	102	4	US-08-974-186-53	Sequence 53, Appl	176	7	1.4	445	4	US-09-642-855-7	Sequence 7, Appl
104	7	1.4	102	4	US-08-795-446B-53	Sequence 2, Appl	177	7	1.4	445	4	US-09-642-855-7	Sequence 7, Appl
105	7	1.4	109	3	US-08-527-04-2	Sequence 2, Appl	178	7	1.4	457	2	US-08-882-704A-6	Sequence 6, Appl
106	7	1.4	109	3	US-09-013-780-2	Sequence 39, Appl	179	7	1.4	457	2	US-09-151-957-6	Sequence 2, Appl
107	7	1.4	111	4	US-09-091-725-39	Sequence 85, Appl	180	7	1.4	483	3	US-09-263-023-2	Sequence 2, Appl
108	7	1.4	113	2	US-08-256-568B-85	Sequence 8, Appl	181	7	1.4	483	3	US-09-471-867-2	Sequence 2, Appl
109	7	1.4	113	4	US-09-038-369B-85	Sequence 20, Appl	182	7	1.4	544	4	US-09-532-594B-18	Sequence 18, Appl
110	7	1.4	127	4	US-09-329-884-8	Sequence 20, Appl	183	7	1.4	548	2	US-08-871-266B-17	Sequence 17, Appl
111	7	1.4	132	1	US-08-611-107-20	Sequence 18, Appl	184	7	1.4	548	2	US-09-018-864A-17	Sequence 17, Appl
112	7	1.4	132	2	US-08-422-560A-20	Sequence 18, Appl	185	7	1.4	548	3	US-08-871-267B-23	Sequence 23, Appl
113	7	1.4	132	4	US-08-468-793-20	Sequence 18, Appl	186	7	1.4	548	3	US-09-618-419-23	Sequence 23, Appl
114	7	1.4	147	4	US-08-751-359-18	Sequence 18, Appl	187	7	1.4	550	4	US-09-387-574-10	Sequence 10, Appl
115	7	1.4	154	3	US-08-907-146-18	Sequence 18, Appl	188	7	1.4	572	3	US-09-668-096-10	Sequence 10, Appl
116	7	1.4	154	3	US-08-486-099-108	Sequence 108, App	189	7	1.4	572	3	US-09-040-681A-4	Sequence 4, Appl
117	7	1.4	154	3	US-08-360-107A-118	Sequence 108, App	190	7	1.4	572	4	US-09-497-897-4	Sequence 4, Appl
118	7	1.4	154	3	US-08-484-223B-108	Sequence 108, App	191	7	1.4	598	4	US-09-532-594B-16	Sequence 16, Appl
119	7	1.4	154	3	US-08-919-597-108	Sequence 108, App	192	7	1.4	604	4	US-08-536-978B-63	Sequence 63, Appl
120	7	1.4	154	3	US-08-475-668A-108	Sequence 108, App	193	7	1.4	606	4	US-08-556-978B-21	Sequence 21, Appl
121	7	1.4	154	3	US-08-485-551A-108	Sequence 108, App	194	7	1.4	606	4	US-08-556-978B-21	Sequence 21, Appl
122	7	1.4	154	3	US-08-471-913A-108	Sequence 108, App	195	7	1.4	606	4	US-09-247-806-4	Sequence 4, Appl
123	7	1.4	154	4	US-08-485-264A-108	Sequence 108, App	196	7	1.4	606	4	US-09-247-806-6	Sequence 6, Appl
124	7	1.4	154	4	US-08-474-349A-108	Sequence 108, App	197	7	1.4	606	4	US-09-247-806-8	Sequence 8, Appl
125	7	1.4	180	6	5273901-7	Patent No. 5273901	198	7	1.4	625	1	US-08-365-981-15	Sequence 15, Appl
126	7	1.4	180	6	5482709-6	Patent No. 5482709	199	7	1.4	643	4	US-09-232-191-33	Sequence 33, Appl
127	7	1.4	227	2	US-08-164-292B-27	Sequence 27, Appl	200	7	1.4	643	4	US-09-232-200-63	Sequence 63, Appl
128	7	1.4	227	3	US-08-845-623-27	Sequence 27, Appl	201	7	1.4	643	4	US-09-232-197-83	Sequence 83, Appl
129	7	1.4	227	3	US-08-815-927-27	Sequence 27, Appl	202	7	1.4	643	4	US-09-232-201-83	Sequence 83, Appl
130	7	1.4	227	4	US-09-103-330-27	Sequence 27, Appl	203	7	1.4	734	4	US-09-532-554B-4	Sequence 4, Appl
131	7	1.4	227	4	US-09-435-242-27	Sequence 27, Appl	204	7	1.4	734	4	US-08-864-038A-3	Sequence 3, Appl
132	7	1.4	235	4	US-09-134-001C-4827	Sequence 4827, Ap	205	7	1.4	804	4	US-09-067-091-2	Sequence 2, Appl
133	7	1.4	241	1	US-07-987-286-17	Sequence 17, Appl	206	7	1.4	805	4	US-09-134-001C-4821	Sequence 4821, Ap
134	7	1.4	241	2	US-08-614-626-17	Sequence 17, Appl	207	7	1.4	924	1	US-08-481-130-28	Sequence 28, Appl
135	7	1.4	243	1	US-07-690-983D-12	Sequence 12, Appl	208	7	1.4	924	1	US-08-656-984A-28	Sequence 28, Appl
136	7	1.4	243	1	US-08-133-804-6	Sequence 6, Appl	209	7	1.4	924	2	US-08-485-604-28	Sequence 28, Appl
137	7	1.4	243	2	US-08-461-386-6	Sequence 6, Appl	210	7	1.4	924	2	US-08-485-595-28	Sequence 28, Appl
138	7	1.4	243	2	US-08-461-386-6	Sequence 6, Appl	211	7	1.4	924	2	US-08-485-595-28	Sequence 28, Appl
139	7	1.4	243	2	US-08-356-786-4	Sequence 2, Appl	212	7	1.4	992	1	US-08-137-499A-1	Sequence 1, Appl
140	7	1.4	247	2	US-08-951-822-23	Sequence 23, Appl	213	7	1.4	1127	4	US-08-462-847-1	Sequence 4, Appl
141	7	1.4	247	3	US-08-705-245-4	Sequence 23, Appl	214	7	1.4	1127	4	US-09-150-460B-11	Sequence 11, Appl
142	7	1.4	247	3	US-09-368-951-23	Sequence 23, Appl	215	7	1.4	1276	4	US-08-937-236-3	Sequence 3, Appl
143	7	1.4	252	4	US-08-462-169B-23	Sequence 23, Appl	216	7	1.4	1291	4	US-08-937-236-6	Sequence 6, Appl
144	7	1.4	252	4	US-08-462-169B-23	Sequence 23, Appl	217	7	1.4	1291	4	US-08-937-236-2	Sequence 2, Appl
145	7	1.4	252	3	US-09-103-079-2	Sequence 9, Appl	218	7	1.4	1292	4	US-08-569-214-5	Sequence 5, Appl
146	7	1.4	252	3	US-09-103-079-2	Sequence 9, Appl	219	7	1.4	1292	4	US-08-569-214-6	Sequence 6, Appl
147	7	1.4	252	3	US-08-851-843A-174	Sequence 23, Appl	220	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
148	7	1.4	261	3	US-08-974-549A-293	Sequence 23, Appl	221	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
149	7	1.4	261	4	US-08-854-050-174	Sequence 174, App	222	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
150	7	1.4	261	4	US-09-430-323-114	Sequence 32, Appl	223	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
151	7	1.4	262	4	US-09-372-422A-32	Sequence 32, Appl	224	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
152	7	1.4	282	4	US-09-455-960-9	Sequence 9, Appl	225	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
153	7	1.4	287	1	US-08-146-421-5	Sequence 5, Appl	226	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
154	7	1.4	302	3	US-09-178-610-6	Sequence 6, Appl	227	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
155	7	1.4	305	4	US-09-540-014-9	Sequence 24, Appl	228	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
156	7	1.4	332	4	US-09-540-014-9	Sequence 15, Appl	229	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
157	7	1.4	351	4	US-08-984-618-15	Sequence 15, Appl	230	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
158	7	1.4	351	4	US-08-984-618-15	Sequence 15, Appl	231	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
159	7	1.4	362	3	US-09-347-975-2	Sequence 1, Appl	232	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
160	7	1.4	374	2	US-07-857-224B-80	Sequence 80, Appl	233	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
161	7	1.4	374	2	US-07-857-224B-81	Sequence 81, Appl	234	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
162	7	1.4	374	2	US-07-857-224B-82	Sequence 82, Appl	235	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
163	7	1.4	374	2	US-07-857-224B-83	Sequence 83, Appl	236	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
164	7	1.4	374	2	US-07-857-224B-84	Sequence 84, Appl	237	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
165	7	1.4	374	2	US-07-857-224B-85	Sequence 85, Appl	238	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
166	7	1.4	376	4	US-09-056-356-202	Sequence 202, App	239	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
167	7	1.4	376	4	US-09-072-596-197	Sequence 197, App	240	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
168	7	1.4	425	3	US-09-109-204-31	Sequence 31, Appl	241	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
169	7	1.4	444	1	US-07-881-075-3	Sequence 3, Appl	242	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
170	7	1.4	444	1	US-08-120-827-3	Sequence 3, Appl	243	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
171	7	1.4	444	1	US-08-478-675-3	Sequence 2, Appl	244	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
172	7	1.4	445	2	US-08-985-090-2	Sequence 2, Appl	245	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl
173	7	1.4	445	3	US-09-165-543-2	Sequence 2, Appl	246	7	1.4	1292	4	US-08-569-214-2	Sequence 2, Appl

247	6	1.2	12	2	US-08-630-645-23	Sequence 23, Appl	320	6	1.2	28	4	US-08-957-001B-25	Sequence 25, Appl
248	6	1.2	12	2	US-08-926-789-7	Sequence 7, Appl	321	6	1.2	28	4	US-09-496-301-25	Sequence 25, Appl
249	6	1.2	12	2	US-08-818-253-41	Sequence 41, Appl	322	6	1.2	28	4	US-08-556-978B-72	Sequence 72, Appl
250	6	1.2	12	3	US-09-166-750-1	Sequence 1, Appl	323	6	1.2	28	4	US-09-288-143-146	Sequence 146, Appl
251	6	1.2	12	3	US-09-166-093-1	Sequence 1, Appl	324	6	1.2	29	1	US-08-244-701B-51	Sequence 51, Appl
252	6	1.2	12	3	US-09-172-019-1	Sequence 1, Appl	325	6	1.2	29	4	US-09-128-440-9	Sequence 9, Appl
253	6	1.2	12	3	US-09-166-094-1	Sequence 41, Appl	326	6	1.2	29	4	US-09-823-494-9	Sequence 9, Appl
254	6	1.2	12	4	US-08-818-252-41	Sequence 3, Appl	327	6	1.2	29	4	US-09-076-721-51	Sequence 51, Appl
255	6	1.2	12	4	US-09-362-805-3	Sequence 3, Appl	328	6	1.2	30	3	US-08-256-747C-44	Sequence 44, Appl
256	6	1.2	12	4	US-09-173-190-3	Sequence 6, Appl	329	6	1.2	30	4	US-08-834-130A-44	Sequence 44, Appl
257	6	1.2	12	4	US-09-069-821-6	Sequence 23, Appl	330	6	1.2	32	1	US-07-828-450-41	Sequence 41, Appl
258	6	1.2	12	4	US-08-766-596A-23	Sequence 57, Appl	331	6	1.2	32	1	US-08-297-330-2	Sequence 2, Appl
259	6	1.2	12	4	US-09-316-919-57	Sequence 7, Appl	332	6	1.2	32	1	US-08-451-240-3	Sequence 2, Appl
260	6	1.2	12	5	PCT-US93-11138-7	Sequence 23, Appl	333	6	1.2	32	1	US-08-470-846A-3	Sequence 3, Appl
261	6	1.2	12	5	PCT-US96-10220-23	Sequence 4, Appl	334	6	1.2	32	2	US-08-954-915A-1	Sequence 1, Appl
262	6	1.2	13	4	US-08-942-456-4	Sequence 66, Appl	335	6	1.2	32	3	US-08-954-915A-48	Sequence 48, Appl
263	6	1.2	13	4	US-08-925-539A-66	Sequence 50, Appl	336	6	1.2	32	3	US-08-942-332B-1	Sequence 1, Appl
264	6	1.2	13	4	US-09-510-721-4	Sequence 8, Appl	337	6	1.2	32	4	US-09-510-721-1	Sequence 1, Appl
265	6	1.2	13	5	PCT-US95-04121-50	Sequence 3, Appl	338	6	1.2	32	4	US-08-924-456-1	Sequence 1, Appl
266	6	1.2	14	1	US-07-961-837-8	Sequence 3, Appl	339	6	1.2	32	4	US-09-510-721-1	Sequence 1, Appl
267	6	1.2	14	1	US-08-556-823-3	Sequence 4, Appl	340	6	1.2	32	4	US-08-924-629C-67	Sequence 67, Appl
268	6	1.2	14	1	US-08-556-823-4	Sequence 3, Appl	341	6	1.2	32	4	US-09-466-268B-6	Sequence 6, Appl
269	6	1.2	14	1	US-08-373-190-3	Sequence 5, Appl	342	6	1.2	32	5	PCT-US94-02391-2	Sequence 2, Appl
270	6	1.2	14	2	US-08-373-190-5	Sequence 3, Appl	343	6	1.2	32	5	PCT-US94-12591-3	Sequence 3, Appl
271	6	1.2	14	2	US-08-438-190A-3	Sequence 5, Appl	344	6	1.2	32	5	PCT-US94-12591-22	Sequence 22, Appl
272	6	1.2	14	2	US-08-438-190A-5	Sequence 3, Appl	345	6	1.2	33	1	US-08-209-747-19	Sequence 19, Appl
273	6	1.2	14	2	US-08-818-253-46	Sequence 46, Appl	346	6	1.2	33	1	US-08-556-823-7	Sequence 7, Appl
274	6	1.2	14	3	US-08-350-213-3	Sequence 3, Appl	347	6	1.2	33	1	US-08-458-298-19	Sequence 19, Appl
275	6	1.2	14	3	US-08-350-213-5	Sequence 5, Appl	348	6	1.2	33	1	US-08-244-701B-1	Sequence 1, Appl
276	6	1.2	14	3	US-09-287-145A-3	Sequence 3, Appl	349	6	1.2	33	1	US-08-244-701B-2	Sequence 2, Appl
277	6	1.2	14	3	US-09-287-145A-5	Sequence 5, Appl	350	6	1.2	33	4	US-09-128-450-8	Sequence 8, Appl
278	6	1.2	14	4	US-09-046-992-6	Sequence 6, Appl	351	6	1.2	33	4	US-09-128-450-17	Sequence 17, Appl
279	6	1.2	14	4	US-08-818-252-46	Sequence 46, Appl	352	6	1.2	33	4	US-09-128-450-23	Sequence 23, Appl
280	6	1.2	14	4	US-09-362-805-8	Sequence 8, Appl	353	6	1.2	33	4	US-09-128-450-29	Sequence 29, Appl
281	6	1.2	14	4	US-09-173-190-8	Sequence 3, Appl	354	6	1.2	33	4	US-09-823-494-8	Sequence 8, Appl
282	6	1.2	14	4	US-09-556-111-3	Sequence 5, Appl	355	6	1.2	33	4	US-09-823-494-17	Sequence 17, Appl
283	6	1.2	14	4	US-09-556-111-5	Sequence 3, Appl	356	6	1.2	33	4	US-08-823-494-23	Sequence 23, Appl
284	6	1.2	14	4	US-08-766-596A-34	Sequence 34, Appl	357	6	1.2	33	4	US-09-823-494-29	Sequence 29, Appl
285	6	1.2	14	4	US-09-316-919-62	Sequence 62, Appl	358	6	1.2	33	4	US-09-076-721-1	Sequence 1, Appl
286	6	1.2	15	2	US-08-528-057-18	Sequence 18, Appl	359	6	1.2	33	4	US-09-076-721-1	Sequence 1, Appl
287	6	1.2	15	2	US-08-373-190-4	Sequence 4, Appl	360	6	1.2	33	4	US-08-766-596A-37	Sequence 37, Appl
288	6	1.2	15	2	US-08-438-190A-4	Sequence 4, Appl	361	6	1.2	34	4	US-08-556-978B-1	Sequence 1, Appl
289	6	1.2	15	3	US-08-350-215-4	Sequence 4, Appl	362	6	1.2	36	1	US-08-209-747-31	Sequence 31, Appl
290	6	1.2	15	3	US-09-287-145A-4	Sequence 4, Appl	363	6	1.2	36	1	US-08-458-298-31	Sequence 31, Appl
291	6	1.2	15	4	US-09-556-111-4	Sequence 4, Appl	364	6	1.2	37	1	US-08-314-596-6	Sequence 6, Appl
292	6	1.2	16	4	US-08-308-494A-2	Sequence 2, Appl	365	6	1.2	37	1	US-08-314-596-22	Sequence 22, Appl
293	6	1.2	16	4	US-08-469-260A-236	Sequence 236, Appl	366	6	1.2	37	1	US-08-320-982-6	Sequence 6, Appl
294	6	1.2	18	1	US-08-240-712-33	Sequence 33, Appl	367	6	1.2	37	1	US-08-320-982-22	Sequence 22, Appl
295	6	1.2	18	1	US-08-443-890-33	Sequence 3, Appl	368	6	1.2	37	3	US-08-819-037-6	Sequence 6, Appl
296	6	1.2	18	3	US-09-053-617A-1	Sequence 1, Appl	369	6	1.2	37	3	US-08-819-037-22	Sequence 22, Appl
297	6	1.2	18	4	US-09-628-588-1	Sequence 5, Appl	370	6	1.2	37	4	US-08-900-230-54	Sequence 54, Appl
298	6	1.2	19	1	US-08-556-823-5	Sequence 5, Appl	371	6	1.2	37	4	US-09-045-940-6	Sequence 6, Appl
299	6	1.2	19	1	US-08-244-701B-3	Sequence 3, Appl	372	6	1.2	37	4	US-09-045-940-22	Sequence 22, Appl
300	6	1.2	19	1	US-08-244-701B-5	Sequence 5, Appl	373	6	1.2	38	1	US-08-318-193-3	Sequence 3, Appl
301	6	1.2	19	4	US-09-076-721-3	Sequence 3, Appl	374	6	1.2	38	6	5200327-14	Patent No. 5200327
302	6	1.2	19	4	US-09-076-721-5	Sequence 5, Appl	375	6	1.2	39	1	US-08-318-193-71	Sequence 71, Appl
303	6	1.2	21	1	US-08-164-151-15	Sequence 15, Appl	376	6	1.2	41	3	US-08-491-954-38	Sequence 38, Appl
304	6	1.2	21	3	US-08-942-332B-4	Sequence 4, Appl	377	6	1.2	41	3	US-08-993-165-9	Sequence 9, Appl
305	6	1.2	21	4	US-09-392-055-12	Sequence 12, Appl	378	6	1.2	41	4	US-09-540-448-9	Sequence 9, Appl
306	6	1.2	21	4	US-08-766-596A-35	Sequence 35, Appl	379	6	1.2	41	4	US-09-466-268B-1	Sequence 1, Appl
307	6	1.2	22	4	US-08-818-148-39	Sequence 39, Appl	380	6	1.2	41	6	5514590-6	Patent No. 5514590
308	6	1.2	23	4	US-09-128-450-7	Sequence 7, Appl	381	6	1.2	42	1	US-08-377-687-39	Sequence 39, Appl
309	6	1.2	23	4	US-08-823-494-7	Sequence 3, Appl	382	6	1.2	42	2	US-08-777-192-39	Sequence 39, Appl
310	6	1.2	24	4	US-08-983-607-33	Sequence 33, Appl	383	6	1.2	42	4	US-08-971-982-39	Sequence 39, Appl
311	6	1.2	25	1	US-08-318-193-74	Sequence 74, Appl	384	6	1.2	43	1	US-08-078-683A-10	Sequence 10, Appl
312	6	1.2	25	4	US-09-466-268B-7	Sequence 7, Appl	385	6	1.2	43	1	US-08-078-683A-11	Sequence 11, Appl
313	6	1.2	26	2	US-08-850-910A-1	Sequence 1, Appl	386	6	1.2	43	3	US-08-178-477B-21	Sequence 21, Appl
314	6	1.2	27	1	US-08-425-069-56	Sequence 56, Appl	387	6	1.2	44	3	US-09-053-197A-55	Sequence 55, Appl
315	6	1.2	27	1	US-08-425-069-59	Sequence 59, Appl	388	6	1.2	44	4	US-09-085-761A-60	Sequence 60, Appl
316	6	1.2	27	2	US-08-317-844B-56	Sequence 56, Appl	389	6	1.2	45	1	US-08-318-193-112	Sequence 12, Appl
317	6	1.2	27	2	US-08-317-844B-59	Sequence 59, Appl	390	6	1.2	45	1	US-08-451-947-97	Sequence 97, Appl
318	6	1.2	28	1	US-08-276-213-4	Sequence 4, Appl	391	6	1.2	45	2	US-08-424-826A-97	Sequence 97, Appl
319	6	1.2	28	2	US-08-640-847C-38	Sequence 38, Appl	392	6	1.2	45	3	US-08-928-694-97	Sequence 97, Appl

393	6	1.2	45	4	US-08-900-230-7	Sequence 7, Appl1	466	6	1.2	111	1	US-08-466-886-43	Sequence 43, Appl1
394	6	1.2	45	5	PCT-US91-06950-97	Sequence 97, Appl1	467	6	1.2	111	4	US-08-466-617-43	Sequence 43, Appl1
395	6	1.2	47	1	US-08-209-747-24	Sequence 24, Appl1	468	6	1.2	112	1	US-08-211-202-133	Sequence 133, App
396	6	1.2	47	1	US-08-209-747-26	Sequence 26, Appl1	469	6	1.2	112	3	US-08-991-813-4	Sequence 4, Appl1
397	6	1.2	47	1	US-08-458-258-24	Sequence 24, Appl1	470	6	1.2	113	2	US-08-256-568B-66	Sequence 86, Appl1
398	6	1.2	47	1	US-08-458-258-26	Sequence 26, Appl1	471	6	1.2	113	2	US-08-458-258-26	Sequence 86, Appl1
399	6	1.2	48	1	US-08-209-747-14	Sequence 14, Appl1	472	6	1.2	113	3	US-08-513-974B-38	Sequence 38, Appl1
400	6	1.2	48	1	US-08-209-747-21	Sequence 21, Appl1	473	6	1.2	115	3	US-09-615-192A-317	Sequence 317, App
401	6	1.2	48	1	US-08-458-298-14	Sequence 14, Appl1	474	6	1.2	117	3	US-08-545-809A-107	Sequence 107, App
402	6	1.2	48	1	US-08-458-298-21	Sequence 21, Appl1	475	6	1.2	117	3	US-09-291-922-12	Sequence 12, Appl1
403	6	1.2	49	1	US-08-209-747-22	Sequence 22, Appl1	476	6	1.2	117	3	US-08-301-162-10	Sequence 10, Appl1
404	6	1.2	49	1	US-08-458-298-22	Sequence 22, Appl1	477	6	1.2	118	3	US-08-545-809A-123	Sequence 123, App
405	6	1.2	50	1	US-08-318-193-24	Sequence 24, Appl1	478	6	1.2	118	4	US-09-461-240-10	Sequence 10, Appl1
406	6	1.2	50	1	US-08-209-747-27	Sequence 27, Appl1	479	6	1.2	118	4	US-09-968-927-10	Sequence 10, Appl1
407	6	1.2	50	1	US-08-458-298-27	Sequence 27, Appl1	480	6	1.2	122	2	US-08-360-125-11	Sequence 11, Appl1
408	6	1.2	50	4	US-08-905-223-488	Sequence 27, Appl1	481	6	1.2	122	2	US-08-450-578-11	Sequence 11, Appl1
409	6	1.2	50	4	US-08-262-856A-1	Sequence 488, App	482	6	1.2	122	2	US-09-017-628-11	Sequence 11, Appl1
410	6	1.2	51	4	US-08-209-747-20	Sequence 1, Appl1	483	6	1.2	122	2	US-09-014-860-11	Sequence 11, Appl1
411	6	1.2	51	1	US-08-209-747-30	Sequence 20, Appl1	484	6	1.2	122	2	US-08-450-363-11	Sequence 11, Appl1
412	6	1.2	51	1	US-08-458-298-30	Sequence 30, Appl1	485	6	1.2	123	1	US-08-078-683A-9	Sequence 9, Appl1
413	6	1.2	51	1	US-08-209-747-6	Sequence 6, Appl1	486	6	1.2	125	4	US-09-160-246-16	Sequence 321, App
414	6	1.2	54	1	US-08-209-747-29	Sequence 29, Appl1	487	6	1.2	127	4	US-09-291-922-4	Sequence 10, Appl1
415	6	1.2	54	1	US-08-458-298-6	Sequence 6, Appl1	488	6	1.2	131	4	US-08-513-974B-325	Sequence 325, App
416	6	1.2	54	1	US-08-209-747-10	Sequence 29, Appl1	489	6	1.2	132	3	US-08-513-974B-326	Sequence 10, Appl1
417	6	1.2	54	1	US-08-905-223-427	Sequence 427, App	490	6	1.2	133	3	US-09-452-229-10	Sequence 45, Appl1
418	6	1.2	55	1	US-08-209-747-10	Sequence 10, Appl1	491	6	1.2	134	2	US-08-850-910A-45	Sequence 48, Appl1
419	6	1.2	55	1	US-08-209-747-32	Sequence 32, Appl1	492	6	1.2	134	2	US-08-850-910A-48	Sequence 13, Appl1
420	6	1.2	55	1	US-08-458-298-10	Sequence 10, Appl1	493	6	1.2	136	2	US-08-505-218-13	Sequence 16, Appl1
421	6	1.2	55	1	US-08-458-298-25	Sequence 25, Appl1	494	6	1.2	136	2	US-08-444-818-16	Sequence 51, Appl1
422	6	1.2	55	1	US-08-458-298-32	Sequence 32, Appl1	495	6	1.2	141	4	US-09-091-725-51	Sequence 2, Appl1
423	6	1.2	55	1	US-08-209-747-28	Sequence 28, Appl1	496	6	1.2	142	1	US-08-556-823-2	Sequence 10, Appl1
424	6	1.2	56	1	US-08-556-823-1	Sequence 1, Appl1	497	6	1.2	142	1	US-08-505-218-2	Sequence 2, Appl1
425	6	1.2	56	1	US-08-458-298-8	Sequence 28, Appl1	498	6	1.2	142	2	US-08-318-193-16	Sequence 16, Appl1
426	6	1.2	56	1	US-08-556-823-9	Sequence 28, Appl1	499	6	1.2	144	2	US-09-372-422A-46	Sequence 46, Appl1
427	6	1.2	56	1	US-08-458-298-28	Sequence 28, Appl1	500	6	1.2	150	4	US-08-318-947A-9	Sequence 9, Appl1
428	6	1.2	56	1	PCT-US96-01720-10	Sequence 10, Appl1	501	6	1.2	151	4	US-08-795-303-9	Sequence 33, Appl1
429	6	1.2	71	5	PCT-US96-01720-11	Sequence 11, Appl1	502	6	1.2	151	4	US-08-933-750C-33	Sequence 33, Appl1
430	6	1.2	71	5	US-09-461-697-42	Sequence 42, Appl1	503	6	1.2	151	2	US-09-234-613-33	Sequence 4, Appl1
431	6	1.2	72	4	US-08-858-207A-368	Sequence 368, App	504	6	1.2	152	2	US-09-096-244-4	Sequence 15, Appl1
432	6	1.2	80	1	US-08-209-747-12	Sequence 12, Appl1	505	6	1.2	153	4	US-08-209-747-15	Sequence 15, Appl1
433	6	1.2	81	4	US-09-134-001C-3940	Sequence 32, Appl1	506	6	1.2	155	1	US-08-458-298-15	Sequence 210, App
434	6	1.2	81	2	US-08-458-298-12	Sequence 86, Appl1	507	6	1.2	155	4	US-09-370-838-210	Sequence 7, Appl1
435	6	1.2	81	2	US-09-014-969-32	Sequence 98, Appl1	508	6	1.2	159	3	US-08-893-165-7	Sequence 196, App
436	6	1.2	85	2	US-08-528-057-8	Sequence 3, Appl1	509	6	1.2	159	4	US-09-540-448-7	Sequence 7, Appl1
437	6	1.2	88	2	US-08-528-057-6	Sequence 3612, Ap	510	6	1.2	159	4	US-09-205-426-196	Sequence 196, App
438	6	1.2	90	4	US-09-247-155-86	Sequence 8, Appl1	511	6	1.2	161	4	US-08-318-193-8	Sequence 318, App
439	6	1.2	90	4	US-08-528-057-6	Sequence 8, Appl1	512	6	1.2	161	4	US-08-615-192A-318	Sequence 16, App
440	6	1.2	90	4	US-09-602-877A-98	Sequence 98, Appl1	513	6	1.2	165	1	US-09-068-195-14	Sequence 14, Appl1
441	6	1.2	91	1	US-08-591-498-3	Sequence 3, Appl1	514	6	1.2	165	4	US-08-195-22	Sequence 22, Appl1
442	6	1.2	91	4	US-08-528-057-5	Sequence 15, Appl1	515	6	1.2	166	4	US-08-858-207A-333	Sequence 333, App
443	6	1.2	92	2	US-08-461-990B-15	Sequence 15, Appl1	516	6	1.2	166	4	US-08-818-112-100	Sequence 100, App
444	6	1.2	92	2	US-08-821-009-1	Sequence 7, Appl1	517	6	1.2	169	2	US-09-072-596-95	Sequence 95, Appl1
445	6	1.2	99	2	US-09-099-572-1	Sequence 1, Appl1	518	6	1.2	169	2	US-08-310-416A-18	Sequence 18, Appl1
446	6	1.2	99	2	US-08-528-057-7	Sequence 15, Appl1	519	6	1.2	169	2	US-08-888-171-18	Sequence 18, Appl1
447	6	1.2	99	2	US-08-821-009-1	Sequence 1, Appl1	520	6	1.2	175	3	US-09-134-001C-9431	Sequence 9431, Ap
448	6	1.2	102	4	US-09-461-697-38	Sequence 83, Appl1	521	6	1.2	177	4	US-09-263-933-25	Sequence 25, Appl1
449	6	1.2	106	4	US-09-199-637A-59	Sequence 38, Appl1	522	6	1.2	181	2	US-08-726-306A-56	Sequence 56, Appl1
450	6	1.2	106	4	US-08-055-006-15	Sequence 3407, Ap	523	6	1.2	182	1	US-08-440-712-29	Sequence 29, Appl1
451	6	1.2	106	4	US-08-055-006-16	Sequence 15, Appl1	524	6	1.2	182	1	US-08-843-890-29	Sequence 29, Appl1
452	6	1.2	106	4	US-08-818-112-83	Sequence 83, Appl1	525	6	1.2	184	2	US-08-693-193-6	Sequence 6, Appl1
453	6	1.2	107	4	US-08-818-112-83	Sequence 83, Appl1	526	6	1.2	184	2	US-09-008-482-2	Sequence 2, Appl1
454	6	1.2	107	4	US-09-056-556-83	Sequence 83, Appl1	527	6	1.2	184	2	US-08-564-090A-6	Sequence 6, Appl1
455	6	1.2	107	4	US-09-072-596-84	Sequence 83, Appl1	528	6	1.2	184	2	PCT-US94-06698-6	Sequence 6, Appl1
456	6	1.2	108	4	US-08-271-252-7	Sequence 5177, Ap	529	6	1.2	186	4	US-09-008-892-6	Sequence 6, Appl1
457	6	1.2	108	4	US-08-379-057-32	Sequence 32, Appl1	530	6	1.2	187	4	US-08-818-112-63	Sequence 63, Appl1
458	6	1.2	109	1	US-08-466-033-91	Sequence 91, Appl1	531	6	1.2	187	4	US-08-818-111-64	Sequence 64, Appl1
459	6	1.2	110	2	US-08-444-733-91	Sequence 91, Appl1	532	6	1.2	187	4		
460	6	1.2	110	2	US-08-464-134-91	Sequence 91, Appl1	533	6	1.2	187	4		
461	6	1.2	110	2	US-08-461-361-91	Sequence 91, Appl1	534	6	1.2	187	4		
462	6	1.2	110	2	US-08-485-910-91	Sequence 91, Appl1	535	6	1.2	187	4		
463	6	1.2	110	2	PCT-US93-06266-75	Sequence 75, Appl1	536	6	1.2	187	4		
464	6	1.2	110	2			537	6	1.2	187	4		
465	6	1.2	110	5			538	6	1.2	187	4		

539	6	1.2	187	4	US-09-056-556-63	Sequence 63, Appl1	612	6	1.2	246	1	US-08-294-386C-3	Sequence 3, Appl1
540	6	1.2	187	4	US-09-072-596-64	Sequence 64, Appl1	613	6	1.2	246	1	US-08-299-249A-1	Sequence 1, Appl1
541	6	1.2	190	2	US-08-466-033-38	Sequence 38, Appl1	614	6	1.2	246	1	US-08-299-249A-10	Sequence 10, Appl1
542	6	1.2	190	2	US-08-444-733-38	Sequence 38, Appl1	615	6	1.2	246	1	US-08-590-708-1	Sequence 1, Appl1
543	6	1.2	190	2	US-08-464-134-38	Sequence 38, Appl1	616	6	1.2	246	5	PCT-US94-06543-1	Sequence 1, Appl1
544	6	1.2	190	2	US-08-461-361-38	Sequence 38, Appl1	617	6	1.2	246	5	PCT-US95-10224-1	Sequence 1, Appl1
545	6	1.2	190	2	US-08-485-910-38	Sequence 38, Appl1	618	6	1.2	246	5	PCT-US95-10224-3	Sequence 3, Appl1
546	6	1.2	190	4	US-09-125-619-27	Sequence 27, Appl1	619	6	1.2	247	1	US-08-324-977-44	Sequence 44, Appl1
547	6	1.2	190	5	PCT-US95-06266-29	Sequence 29, Appl1	620	6	1.2	247	2	US-08-384-616-44	Sequence 44, Appl1
548	6	1.2	194	3	US-08-480-640A-196	Sequence 196, App	621	6	1.2	247	2	US-08-304-686A-44	Sequence 44, Appl1
549	6	1.2	194	4	US-08-686-968C-196	Sequence 196, App	622	6	1.2	247	4	US-09-315-850-44	Sequence 44, Appl1
550	6	1.2	194	4	US-08-488-237A-196	Sequence 196, App	623	6	1.2	250	4	US-09-134-001C-2869	Sequence 2869, Ap
551	6	1.2	194	4	US-08-375-992A-196	Sequence 196, App	624	6	1.2	251	1	US-08-209-747-8	Sequence 8, Appl1
552	6	1.2	195	4	US-09-288-143-144	Sequence 144, App	625	6	1.2	251	1	US-08-458-298-8	Sequence 8, Appl1
553	6	1.2	195	4	US-09-125-619-17	Sequence 17, Appl1	626	6	1.2	253	1	US-08-242-188-2	Sequence 2, Appl1
554	6	1.2	197	4	US-09-125-619-25	Sequence 25, Appl1	627	6	1.2	253	1	US-08-509-261A-2	Sequence 2, Appl1
555	6	1.2	198	1	US-08-278-091-16	Sequence 16, Appl1	628	6	1.2	253	1	US-08-660-626-8	Sequence 8, Appl1
556	6	1.2	198	1	US-08-483-859-16	Sequence 16, Appl1	629	6	1.2	253	1	US-08-692-892-2	Sequence 2, Appl1
557	6	1.2	198	1	US-08-472-173-16	Sequence 16, Appl1	630	6	1.2	253	2	US-08-713-939A-2	Sequence 2, Appl1
558	6	1.2	198	2	US-08-487-167-16	Sequence 16, Appl1	631	6	1.2	253	2	US-08-668-162A-22	Sequence 22, Appl1
559	6	1.2	198	2	US-08-482-816-16	Sequence 16, Appl1	632	6	1.2	253	4	US-09-031-168-8	Sequence 8, Appl1
560	6	1.2	198	2	US-08-296-149-16	Sequence 16, Appl1	633	6	1.2	253	4	US-09-128-450-20	Sequence 20, Appl1
561	6	1.2	198	2	US-08-801-499-16	Sequence 16, Appl1	634	6	1.2	253	4	US-09-036-579-2	Sequence 2, Appl1
562	6	1.2	198	3	US-08-615-271-16	Sequence 16, Appl1	635	6	1.2	253	4	US-09-823-494-20	Sequence 20, Appl1
563	6	1.2	198	3	US-09-074-660-16	Sequence 16, Appl1	636	6	1.2	253	4	US-09-550-374-2	Sequence 2, Appl1
564	6	1.2	198	3	US-09-074-659-16	Sequence 16, Appl1	637	6	1.2	254	1	US-08-242-188-1	Sequence 1, Appl1
565	6	1.2	198	3	US-09-106-468-16	Sequence 16, Appl1	638	6	1.2	254	1	US-08-236-918A-4	Sequence 4, Appl1
566	6	1.2	198	4	US-09-106-466A-16	Sequence 16, Appl1	639	6	1.2	254	1	US-08-509-261A-1	Sequence 1, Appl1
567	6	1.2	198	4	US-09-106-467-16	Sequence 16, Appl1	640	6	1.2	254	1	US-08-660-626-7	Sequence 7, Appl1
568	6	1.2	200	4	US-08-965-056-104	Sequence 104, App	641	6	1.2	254	1	US-08-692-892-1	Sequence 1, Appl1
569	6	1.2	200	4	US-09-562-737-107	Sequence 107, App	642	6	1.2	254	2	US-08-713-939A-1	Sequence 1, Appl1
570	6	1.2	201	4	US-09-134-001C-4319	Sequence 4319, Ap	643	6	1.2	254	2	US-08-668-162A-21	Sequence 21, Appl1
571	6	1.2	202	4	US-09-199-637A-149	Sequence 149, App	644	6	1.2	254	4	US-09-031-168-7	Sequence 7, Appl1
572	6	1.2	208	4	US-09-128-450-18	Sequence 18, Appl1	645	6	1.2	254	4	US-09-128-450-19	Sequence 19, Appl1
573	6	1.2	208	4	US-09-823-494-18	Sequence 18, Appl1	646	6	1.2	254	4	US-09-128-450-26	Sequence 26, Appl1
574	6	1.2	209	4	US-09-134-001C-3637	Sequence 3637, Ap	647	6	1.2	254	4	US-09-128-450-28	Sequence 28, Appl1
575	6	1.2	213	3	US-08-971-158-4	Sequence 4, Appl1	648	6	1.2	254	4	US-09-036-579-1	Sequence 1, Appl1
576	6	1.2	213	4	US-09-416-488-4	Sequence 4, Appl1	649	6	1.2	254	4	US-09-823-494-19	Sequence 19, Appl1
577	6	1.2	214	2	US-08-984-172-1	Sequence 30, Appl1	650	6	1.2	254	4	US-09-823-494-26	Sequence 26, Appl1
578	6	1.2	214	4	US-08-861-774E-30	Sequence 30, Appl1	651	6	1.2	254	4	US-09-823-494-28	Sequence 28, Appl1
579	6	1.2	215	4	US-09-134-001C-4736	Sequence 4736, Ap	652	6	1.2	254	4	US-09-150-864A-4	Sequence 4, Appl1
580	6	1.2	216	4	US-09-125-619-46	Sequence 46, Appl1	653	6	1.2	254	4	US-09-550-374-1	Sequence 1, Appl1
581	6	1.2	217	4	US-09-416-509C-1	Sequence 1, Appl1	654	6	1.2	255	1	US-08-242-188-4	Sequence 4, Appl1
582	6	1.2	219	1	US-08-463-115-91	Sequence 91, Appl1	655	6	1.2	255	1	US-08-509-261A-4	Sequence 4, Appl1
583	6	1.2	219	1	US-08-465-388-91	Sequence 91, Appl1	656	6	1.2	255	1	US-08-660-626-10	Sequence 10, Appl1
584	6	1.2	222	3	US-08-924-747-22	Sequence 22, Appl1	657	6	1.2	255	1	US-08-692-892-4	Sequence 4, Appl1
585	6	1.2	222	4	US-09-247-373B-22	Sequence 22, Appl1	658	6	1.2	255	2	US-08-713-939A-4	Sequence 4, Appl1
586	6	1.2	222	4	US-09-296-715-22	Sequence 22, Appl1	659	6	1.2	255	2	US-08-668-162A-24	Sequence 24, Appl1
587	6	1.2	223	5	PCT-US94-14277-6	Sequence 6, Appl1	660	6	1.2	255	4	US-09-031-168-10	Sequence 10, Appl1
588	6	1.2	225	4	US-09-134-001C-3215	Sequence 3215, Ap	661	6	1.2	255	4	US-09-036-579-4	Sequence 4, Appl1
589	6	1.2	230	1	US-08-118-469A-3	Sequence 3, Appl1	662	6	1.2	255	4	US-09-550-374-4	Sequence 4, Appl1
590	6	1.2	230	1	US-08-809-119-3	Sequence 3, Appl1	663	6	1.2	256	4	US-09-128-450-22	Sequence 22, Appl1
591	6	1.2	230	2	US-08-928-692-30	Sequence 30, Appl1	664	6	1.2	256	4	US-09-823-494-22	Sequence 22, Appl1
592	6	1.2	230	4	US-08-768-373-4	Sequence 4, Appl1	665	6	1.2	258	4	US-08-957-130-13	Sequence 13, Appl1
593	6	1.2	230	4	US-08-818-112-65	Sequence 65, Appl1	666	6	1.2	260	4	US-09-452-2339-12	Sequence 12, Appl1
594	6	1.2	230	4	US-09-339-972-30	Sequence 30, Appl1	667	6	1.2	262	4	US-08-154-364-41	Sequence 41, Appl1
595	6	1.2	230	4	US-08-818-111-66	Sequence 66, Appl1	668	6	1.2	263	1	US-07-721-761A-51	Sequence 51, Appl1
596	6	1.2	230	4	US-09-056-556-65	Sequence 65, Appl1	669	6	1.2	263	1	US-08-242-188-3	Sequence 3, Appl1
597	6	1.2	230	4	US-09-072-596-66	Sequence 66, Appl1	670	6	1.2	263	1	US-08-509-261A-3	Sequence 3, Appl1
598	6	1.2	233	2	US-08-928-692-31	Sequence 31, Appl1	671	6	1.2	263	1	US-08-660-626-9	Sequence 9, Appl1
599	6	1.2	233	4	US-09-339-972-31	Sequence 31, Appl1	672	6	1.2	263	1	US-08-692-892-3	Sequence 3, Appl1
600	6	1.2	237	3	US-09-216-001-1	Sequence 1, Appl1	673	6	1.2	263	2	US-08-713-939A-3	Sequence 3, Appl1
601	6	1.2	237	4	US-08-878-862-1	Sequence 1, Appl1	674	6	1.2	263	2	US-08-668-162A-23	Sequence 23, Appl1
602	6	1.2	239	6	5455030-13	Patent No. 5455030	675	6	1.2	263	4	US-09-031-168-9	Sequence 9, Appl1
603	6	1.2	240	3	US-08-942-001-9	Sequence 9, Appl1	676	6	1.2	263	4	US-09-336-579-3	Sequence 3, Appl1
604	6	1.2	240	4	US-09-337-386-9	Sequence 9, Appl1	677	6	1.2	263	4	US-09-550-374-3	Sequence 3, Appl1
605	6	1.2	240	4	US-09-846-922-9	Sequence 9, Appl1	678	6	1.2	264	4	US-09-128-450-21	Sequence 21, Appl1
606	6	1.2	241	2	US-08-700-607-3	Sequence 3, Appl1	679	6	1.2	264	4	US-09-823-494-21	Sequence 21, Appl1
607	6	1.2	242	6	5455030-17	Patent No. 5455030	680	6	1.2	265	1	US-08-468-763-19	Sequence 19, Appl1
608	6	1.2	243	4	US-09-216-295-13	Sequence 13, Appl1	681	6	1.2	265	2	US-08-393-996A-19	Sequence 19, Appl1
609	6	1.2	245	4	US-08-918-148-78	Sequence 78, Appl1	682	6	1.2	266	4	US-09-615-192A-394	Sequence 394, App
610	6	1.2	246	1	US-08-286-888B-1	Sequence 1, Appl1	683	6	1.2	266	6	5223391-7	Patent No. 5223391
611	6	1.2	246	1	US-08-294-386C-1	Sequence 1, Appl1	684	6	1.2	267	1	US-07-706-872-3	Sequence 3, Appl1

685	6	1.2	268	3	US-08-871-483-11	Sequence 11, Appl	758	6	1.2	342	4	US-09-129-033-2	Sequence 2, Appl
686	6	1.2	269	3	US-09-100-557-1	Sequence 1, Appl	759	6	1.2	344	1	US-07-941-523-24	Sequence 24, Appl
687	6	1.2	269	3	US-08-926-842B-59	Sequence 159, Appl	760	6	1.2	344	1	US-07-857-224B-93	Sequence 93, Appl
688	6	1.2	270	4	US-09-134-001C-3167	Sequence 3167, Ap	761	6	1.2	344	2	US-08-602-725-34	Sequence 34, Appl
689	6	1.2	273	4	US-09-008-882-2	Sequence 2, Appl	762	6	1.2	344	2	US-09-134-001C-3158	Sequence 3158, Ap
690	6	1.2	277	4	US-09-134-001C-5261	Sequence 5261, Ap	763	6	1.2	344	4	US-09-134-001C-4304	Sequence 4304, Ap
691	6	1.2	278	1	US-07-941-651-2	Sequence 2, Appl	764	6	1.2	345	1	US-08-171-382-6	Sequence 6, Appl
692	6	1.2	278	1	US-07-988-260B-3	Sequence 3, Appl	765	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
693	6	1.2	278	1	US-08-188-582-9	Sequence 9, Appl	766	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
694	6	1.2	278	1	US-08-279-986-2	Sequence 2, Appl	767	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
695	6	1.2	278	1	US-08-646-715-9	Sequence 9, Appl	768	6	1.2	345	5	US-09-294-531B-31	Sequence 31, Appl
696	6	1.2	278	1	US-08-646-715-9	Sequence 9, Appl	769	6	1.2	345	5	US-09-294-531B-31	Sequence 31, Appl
697	6	1.2	278	5	PCT-US96-03916-60	Sequence 60, Appl	770	6	1.2	347	3	PCT-US95-11878-6	Sequence 6, Appl
698	6	1.2	278	5	PCT-US96-03916-72	Sequence 72, Appl	771	6	1.2	347	3	PCT-US95-11878-6	Sequence 6, Appl
699	6	1.2	280	1	US-08-434-255-8	Sequence 8, Appl	772	6	1.2	347	3	US-09-144-914-6	Sequence 6, Appl
700	6	1.2	280	1	US-08-459-967-8	Sequence 8, Appl	773	6	1.2	348	4	US-09-216-293-16	Sequence 6, Appl
701	6	1.2	280	1	US-08-460-327-8	Sequence 8, Appl	774	6	1.2	349	4	US-09-216-293-16	Sequence 6, Appl
702	6	1.2	280	1	US-08-459-871-8	Sequence 8, Appl	775	6	1.2	351	1	US-08-471-770-46	Sequence 8, Appl
703	6	1.2	280	4	US-09-160-246-14	Sequence 14, Appl	776	6	1.2	351	2	US-08-468-059-46	Sequence 46, Appl
704	6	1.2	280	4	US-09-024-532-2	Sequence 2, Appl	777	6	1.2	351	2	US-08-468-059-46	Sequence 46, Appl
705	6	1.2	280	4	US-09-104-623A-2	Sequence 2, Appl	778	6	1.2	351	4	US-09-109-910-46	Sequence 46, Appl
706	6	1.2	280	4	US-09-019-532-2	Sequence 2, Appl	779	6	1.2	353	4	US-09-134-001C-4640	Sequence 46, Appl
707	6	1.2	283	2	US-08-852-401-4	Sequence 2, Appl	780	6	1.2	353	4	US-08-153-848-28	Sequence 28, Appl
708	6	1.2	285	2	US-08-852-401-4	Sequence 2, Appl	781	6	1.2	355	1	US-08-153-848-28	Sequence 28, Appl
709	6	1.2	285	2	US-08-852-401-4	Sequence 2, Appl	782	6	1.2	355	3	US-09-299-843A-32	Sequence 32, Appl
710	6	1.2	293	1	US-08-310-416A-16	Sequence 16, Appl	783	6	1.2	355	3	US-09-299-843A-32	Sequence 32, Appl
711	6	1.2	293	1	US-08-868-171-16	Sequence 16, Appl	784	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
712	6	1.2	296	1	US-07-783-705A-5	Sequence 5, Appl	785	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
713	6	1.2	298	3	US-08-961-871-10	Sequence 10, Appl	786	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
714	6	1.2	299	6	5514590-4	Patent No. 5514590	787	6	1.2	355	5	PCT-US93-11153-32	Sequence 32, Appl
715	6	1.2	308	2	US-09-047-026A-23	Sequence 23, Appl	788	6	1.2	356	2	US-08-461-990B-2	Sequence 2, Appl
716	6	1.2	309	2	US-08-405-175A-7	Sequence 7, Appl	789	6	1.2	357	4	US-08-638-911A-37	Sequence 37, Appl
717	6	1.2	309	2	US-08-405-175A-8	Sequence 8, Appl	790	6	1.2	357	4	US-08-638-911A-37	Sequence 37, Appl
718	6	1.2	309	2	US-08-508-761B-32	Sequence 32, Appl	791	6	1.2	357	4	US-08-638-911A-37	Sequence 37, Appl
719	6	1.2	310	3	US-08-897-097-4	Sequence 4, Appl	792	6	1.2	360	3	US-08-586-165-2	Sequence 2, Appl
720	6	1.2	311	3	US-08-897-097-4	Sequence 3, Appl	793	6	1.2	360	4	US-08-899-437-7	Sequence 7, Appl
721	6	1.2	311	3	US-08-897-097-4	Sequence 3, Appl	794	6	1.2	360	4	US-09-126-121-7	Sequence 7, Appl
722	6	1.2	313	4	US-09-336-643A-81	Sequence 81, Appl	795	6	1.2	362	3	US-08-850-338-4	Sequence 4, Appl
723	6	1.2	315	4	US-09-336-643A-81	Sequence 81, Appl	796	6	1.2	362	3	US-08-850-338-4	Sequence 4, Appl
724	6	1.2	317	2	US-08-864-799-4	Sequence 2, Appl	797	6	1.2	363	4	US-09-126-121-3	Sequence 3, Appl
725	6	1.2	317	2	US-08-864-799-5	Sequence 2, Appl	798	6	1.2	363	4	US-09-301-665-4	Sequence 4, Appl
726	6	1.2	317	2	US-08-864-799-5	Sequence 2, Appl	799	6	1.2	370	1	US-08-434-225-6	Sequence 6, Appl
727	6	1.2	318	4	US-08-937-067-6	Sequence 6, Appl	800	6	1.2	370	1	US-08-459-967-6	Sequence 6, Appl
728	6	1.2	318	4	US-08-278-042C-3	Sequence 3, Appl	801	6	1.2	370	1	US-08-459-967-6	Sequence 6, Appl
729	6	1.2	321	2	US-09-060-756-727	Sequence 727, App	802	6	1.2	370	2	US-08-459-967-6	Sequence 6, Appl
730	6	1.2	321	6	US-08-937-972-3	Sequence 3, Appl	803	6	1.2	370	2	US-08-528-057-12	Sequence 12, Appl
731	6	1.2	323	3	5169835-17	Patent No. 5169835	804	6	1.2	370	4	US-09-199-737-2	Sequence 2, Appl
732	6	1.2	323	3	US-08-853-839-2	Sequence 2, Appl	805	6	1.2	370	4	US-08-900-220-2	Sequence 2, Appl
733	6	1.2	323	3	US-09-041-889-28	Sequence 28, Appl	806	6	1.2	370	4	US-09-058-333A-2	Sequence 2, Appl
734	6	1.2	325	2	US-09-615-192A-395	Sequence 395, App	807	6	1.2	373	2	US-09-724-224-2	Sequence 2, Appl
735	6	1.2	325	2	US-08-107-676-30	Sequence 30, Appl	808	6	1.2	373	2	US-08-528-057-14	Sequence 14, Appl
736	6	1.2	325	3	US-08-303-861-12	Sequence 12, Appl	809	6	1.2	374	1	US-08-625-876-2	Sequence 2, Appl
737	6	1.2	325	4	US-08-975-762-8	Sequence 8, Appl	810	6	1.2	374	1	US-08-625-876-2	Sequence 2, Appl
738	6	1.2	325	4	US-08-821-324-8	Sequence 8, Appl	811	6	1.2	375	2	US-08-625-876-6	Sequence 6, Appl
739	6	1.2	325	4	US-09-295-028-8	Sequence 8, Appl	812	6	1.2	375	2	US-07-857-224B-86	Sequence 86, Appl
740	6	1.2	327	3	US-09-106-582-8	Sequence 8, Appl	813	6	1.2	376	4	US-07-857-224B-88	Sequence 88, Appl
741	6	1.2	328	3	US-08-513-974B-372	Sequence 372, App	814	6	1.2	376	4	US-07-857-224B-88	Sequence 88, Appl
742	6	1.2	328	3	US-08-459-046-2	Sequence 2, Appl	815	6	1.2	377	3	US-08-528-057-2	Sequence 2, Appl
743	6	1.2	328	3	US-08-513-974B-39	Sequence 39, Appl	816	6	1.2	377	3	US-09-041-889-29	Sequence 29, Appl
744	6	1.2	328	3	US-08-513-974B-56	Sequence 56, Appl	817	6	1.2	378	2	US-08-528-057-2	Sequence 2, Appl
745	6	1.2	328	3	US-08-513-974B-371	Sequence 371, App	818	6	1.2	378	2	US-08-528-057-2	Sequence 2, Appl
746	6	1.2	329	4	US-08-513-974B-380	Sequence 380, App	819	6	1.2	378	5	US-08-244-205-7	Sequence 5, Appl
747	6	1.2	330	4	US-08-690-347-2	Sequence 2, Appl	820	6	1.2	378	5	US-09-171-461-6	Sequence 6, Appl
748	6	1.2	331	2	US-08-907-674-1	Sequence 1, Appl	821	6	1.2	380	4	PCT-US92-10284-7	Sequence 7, Appl
749	6	1.2	331	2	US-08-907-674-1	Sequence 1, Appl	822	6	1.2	380	4	US-08-904-259A-8	Sequence 8, Appl
750	6	1.2	331	2	US-09-391-959-1	Sequence 1, Appl	823	6	1.2	380	4	US-09-222-594-8	Sequence 8, Appl
751	6	1.2	332	5	US-08-405-175A-5	Sequence 5, Appl	824	6	1.2	382	4	US-08-444-818-68	Sequence 68, Appl
752	6	1.2	333	4	PCT-US94-14277-2	Sequence 2, Appl	825	6	1.2	382	4	US-08-969-815-2	Sequence 2, Appl
753	6	1.2	333	4	US-09-556-877-292	Sequence 292, App	826	6	1.2	382	4	US-09-120-025-2	Sequence 2, Appl
754	6	1.2	334	4	US-09-620-412C-292	Sequence 292, App	827	6	1.2	384	2	US-09-710-481-2	Sequence 2, Appl
755	6	1.2	336	4	US-09-060-756-728	Sequence 728, App	828	6	1.2	384	3	US-08-401-068-2	Sequence 2, Appl
756	6	1.2	337	2	US-09-134-001C-5297	Sequence 5297, Ap	829	6	1.2	384	4	US-08-846-338-8	Sequence 8, Appl
757	6	1.2	341	2	US-09-013-634-2	Sequence 7, Appl	830	6	1.2	384	6	US-08-411-768B-7	Sequence 7, Appl
					US-08-209-521-11	Sequence 11, Appl						US-08-139-195-2	Sequence 2, Appl
												5514787-2	Patent No. 5514787
												US-08-605-106-13	Sequence 13, Appl

831	6	1.2	391	1	US-08-602-010A-6	Sequence 6, Appl1	904	6	1.2	458	1	US-08-336-618-24	Sequence 24, Appl1
832	6	1.2	391	1	US-08-680-726A-6	Sequence 6, Appl1	905	6	1.2	460	4	US-08-444-818-20	Sequence 20, Appl1
833	6	1.2	391	4	US-09-092-409-6	Sequence 6, Appl1	906	6	1.2	463	4	US-08-845-258-25	Sequence 25, Appl1
834	6	1.2	394	2	US-08-839-581A-4	Sequence 4, Appl1	907	6	1.2	463	4	US-08-990-571-25	Sequence 25, Appl1
835	6	1.2	394	4	US-09-023-591A-4	Sequence 4, Appl1	908	6	1.2	463	4	US-08-723-142A-25	Sequence 25, Appl1
836	6	1.2	397	1	US-07-956-697B-5	Sequence 5, Appl1	909	6	1.2	463	4	US-09-528-784A-25	Sequence 25, Appl1
837	6	1.2	397	1	US-08-263-098-5	Sequence 5, Appl1	910	6	1.2	464	2	US-08-602-725-32	Sequence 32, Appl1
838	6	1.2	397	1	US-08-434-258-2	Sequence 2, Appl1	911	6	1.2	464	4	US-08-970-134-3	Sequence 3, Appl1
839	6	1.2	397	1	US-08-434-255-4	Sequence 4, Appl1	912	6	1.2	465	4	US-08-090-369-1	Sequence 1, Appl1
840	6	1.2	397	1	US-08-459-871-4	Sequence 2, Appl1	913	6	1.2	465	4	US-09-482-971-1	Sequence 1, Appl1
841	6	1.2	397	1	US-08-459-867-2	Sequence 4, Appl1	914	6	1.2	465	4	US-09-134-001C-3856	Sequence 3856, Ap
842	6	1.2	397	1	US-08-460-327-2	Sequence 2, Appl1	915	6	1.2	466	4	US-08-970-134-1	Sequence 1, Appl1
843	6	1.2	397	1	US-08-460-327-2	Sequence 4, Appl1	916	6	1.2	466	4	US-08-914-375C-60	Sequence 60, Appl1
844	6	1.2	397	1	US-08-459-871-2	Sequence 2, Appl1	917	6	1.2	471	3	US-09-106-164-2	Sequence 2, Appl1
845	6	1.2	397	1	US-08-459-871-4	Sequence 4, Appl1	918	6	1.2	475	1	US-07-115-184-4	Sequence 4, Appl1
846	6	1.2	397	4	US-09-459-133-2	Sequence 2, Appl1	919	6	1.2	475	1	US-07-876-280-7	Sequence 7, Appl1
847	6	1.2	400	4	US-09-086-010-2	Sequence 2, Appl1	920	6	1.2	475	1	US-07-876-280-28	Sequence 28, Appl1
848	6	1.2	401	2	US-08-278-042C-2	Sequence 2, Appl1	921	6	1.2	475	1	US-07-935-310A-2	Sequence 2, Appl1
849	6	1.2	403	3	US-08-776-246-4	Sequence 4, Appl1	922	6	1.2	475	1	US-07-828-788A-2	Sequence 2, Appl1
850	6	1.2	404	2	US-08-244-205-9	Sequence 9, Appl1	923	6	1.2	475	1	US-08-049-783-4	Sequence 4, Appl1
851	6	1.2	404	5	PCT-US92-10284-9	Sequence 9, Appl1	924	6	1.2	475	1	US-08-147-189-2	Sequence 2, Appl1
852	6	1.2	405	4	US-09-144-914-5	Sequence 5, Appl1	925	6	1.2	475	1	US-08-316-301A-8	Sequence 8, Appl1
853	6	1.2	407	4	US-09-271-438A-4	Sequence 4, Appl1	926	6	1.2	475	2	US-08-904-278-2	Sequence 2, Appl1
854	6	1.2	407	4	US-09-271-438A-10	Sequence 10, Appl1	927	6	1.2	475	4	US-08-904-278-4	Sequence 4, Appl1
855	6	1.2	410	4	US-09-134-001C-4135	Sequence 4135, Ap	928	6	1.2	475	4	US-09-076-137-8	Sequence 8, Appl1
856	6	1.2	411	2	US-08-773-870-1	Sequence 1, Appl1	929	6	1.2	475	4	US-07-941-650A-6	Sequence 6, Appl1
857	6	1.2	412	4	US-09-423-439-10	Sequence 10, Appl1	930	6	1.2	475	4	US-09-222-594-2	Sequence 2, Appl1
858	6	1.2	413	1	US-08-295-676A-2	Sequence 2, Appl1	931	6	1.2	475	4	US-09-222-594-4	Sequence 4, Appl1
859	6	1.2	413	1	US-08-750-077-6	Sequence 6, Appl1	932	6	1.2	475	5	PCT-US92-03624-8	Sequence 8, Appl1
860	6	1.2	413	2	US-08-481-814A-8	Sequence 8, Appl1	933	6	1.2	475	5	PCT-US92-11337-2	Sequence 2, Appl1
861	6	1.2	413	3	US-08-948-591-2	Sequence 2, Appl1	934	6	1.2	475	5	PCT-US93-07409-2	Sequence 2, Appl1
862	6	1.2	413	4	US-09-265-566-2	Sequence 2, Appl1	935	6	1.2	485	1	US-07-881-075-1	Sequence 1, Appl1
863	6	1.2	413	4	US-09-242-737-4	Sequence 2, Appl1	936	6	1.2	485	1	US-08-120-827-1	Sequence 1, Appl1
864	6	1.2	416	4	US-09-330-611-8	Sequence 4, Appl1	937	6	1.2	485	1	US-08-478-612-2	Sequence 1, Appl1
865	6	1.2	421	4	US-09-332-478-6	Sequence 6, Appl1	938	6	1.2	485	4	US-09-384-275-2	Sequence 4, Appl1
866	6	1.2	422	4	US-08-712-072C-2	Sequence 2, Appl1	939	6	1.2	491	4	US-09-262-856A-5	Sequence 5, Appl1
867	6	1.2	424	4	US-09-120-817-2	Sequence 2, Appl1	940	6	1.2	491	4	US-09-134-001C-4612	Sequence 4612, Ap
868	6	1.2	427	4	US-09-199-737-4	Sequence 4, Appl1	941	6	1.2	492	4	US-08-134-001C-4847	Sequence 4847, Ap
869	6	1.2	427	4	US-08-900-230-4	Sequence 4, Appl1	942	6	1.2	493	1	US-08-362-512A-4	Sequence 4, Appl1
870	6	1.2	427	4	US-09-058-333A-4	Sequence 4, Appl1	943	6	1.2	494	3	US-08-964-939-4	Sequence 4, Appl1
871	6	1.2	428	2	US-08-717-587A-2	Sequence 2, Appl1	944	6	1.2	496	4	US-09-134-001C-3703	Sequence 3703, Ap
872	6	1.2	428	3	US-08-883-610A-2	Sequence 3, Appl1	945	6	1.2	497	4	US-09-594-193-2	Sequence 2, Appl1
873	6	1.2	428	3	US-09-118-319-5	Sequence 5, Appl1	946	6	1.2	498	4	US-09-232-468A-18	Sequence 18, Appl1
874	6	1.2	428	4	US-08-936-094A-2	Sequence 2, Appl1	947	6	1.2	498	4	US-08-222-468A-24	Sequence 24, Appl1
875	6	1.2	430	5	PCT-US93-04392-3	Sequence 6, Appl1	948	6	1.2	498	4	US-08-686-968C-231	Sequence 231, App
876	6	1.2	432	2	US-08-904-278-6	Sequence 2, Appl1	949	6	1.2	498	4	US-09-103-754A-5	Sequence 5, Appl1
877	6	1.2	432	4	US-09-222-594-6	Sequence 6, Appl1	950	6	1.2	500	4	US-09-326-203A-17	Sequence 17, Appl1
878	6	1.2	433	1	US-07-661-610C-2	Sequence 2, Appl1	951	6	1.2	505	3	US-08-303-861-2	Sequence 2, Appl1
879	6	1.2	433	4	US-09-364-230-14	Sequence 14, Appl1	952	6	1.2	507	4	US-09-091-097-10	Sequence 10, Appl1
880	6	1.2	433	4	US-09-134-001C-3085	Sequence 3085, Ap	953	6	1.2	507	4	US-08-724-281-1	Sequence 1, Appl1
881	6	1.2	434	2	US-08-710-249-4	Sequence 4, Appl1	954	6	1.2	508	2	US-08-464-266-2	Sequence 2, Appl1
882	6	1.2	434	4	US-09-220-157A-4	Sequence 4, Appl1	955	6	1.2	512	3	US-09-011-730-2	Sequence 3, Appl1
883	6	1.2	437	4	US-09-134-001C-3418	Sequence 3418, Ap	956	6	1.2	512	3	US-09-015-296-3	Sequence 3, Appl1
884	6	1.2	437	4	US-09-052-778-2	Sequence 2, Appl1	957	6	1.2	512	4	US-08-867-611-58	Sequence 58, Appl1
885	6	1.2	442	1	US-08-399-986B-2	Sequence 2, Appl1	958	6	1.2	512	4	US-08-724-224-4	Sequence 4, Appl1
886	6	1.2	443	2	US-08-493-754A-2	Sequence 2, Appl1	959	6	1.2	512	4	US-09-533-722-3	Sequence 3, Appl1
887	6	1.2	443	4	US-09-252-292C-27	Sequence 27, Appl1	960	6	1.2	513	1	US-08-464-266-2	Sequence 2, Appl1
888	6	1.2	444	5	PCT-US94-05387-7	Sequence 7, Appl1	961	6	1.2	513	1	US-08-200-232-4	Sequence 4, Appl1
889	6	1.2	445	5	US-08-244-205-5	Sequence 5, Appl1	962	6	1.2	513	2	US-08-459-847-8	Sequence 8, Appl1
890	6	1.2	446	4	US-09-161-994A-5	Sequence 5, Appl1	963	6	1.2	513	3	US-08-889-419-19	Sequence 19, Appl1
891	6	1.2	446	5	PCT-US92-10284-5	Sequence 5, Appl1	964	6	1.2	513	3	US-08-464-266-2	Sequence 2, Appl1
892	6	1.2	446	5	PCT-US94-01321-10	Sequence 10, Appl1	965	6	1.2	513	4	US-08-464-514-2	Sequence 2, Appl1
893	6	1.2	446	5	US-08-861-464-2	Sequence 2, Appl1	966	6	1.2	513	4	US-08-486-403-2	Sequence 2, Appl1
894	6	1.2	450	2	US-08-396-001-2	Sequence 2, Appl1	967	6	1.2	513	4	US-08-402-542-19	Sequence 19, Appl1
895	6	1.2	450	2	US-09-323-433A-2	Sequence 2, Appl1	968	6	1.2	513	5	PCT-US93-07189-19	Sequence 19, Appl1
896	6	1.2	450	4	US-08-171-382-4	Sequence 4, Appl1	969	6	1.2	513	5	PCT-US95-02219-4	Sequence 4, Appl1
897	6	1.2	454	1	US-08-309-420-4	Sequence 4, Appl1	970	6	1.2	513	5	Patent No. 5187076	Sequence 13, Appl1
898	6	1.2	454	1	US-09-294-531B-30	Sequence 30, Appl1	971	6	1.2	514	5	Sequence 13, Appl1	
899	6	1.2	454	5	PCT-US95-11856-4	Sequence 4, Appl1	972	6	1.2	514	5		
900	6	1.2	454	5	PCT-US95-11878-4	Sequence 4, Appl1	973	6	1.2	514	5		
901	6	1.2	454	5			974	6	1.2	514	5		
902	6	1.2	454	5			975	6	1.2	514	5		
903	6	1.2	454	5			976	6	1.2	514	5		



```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106009/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106010/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106011/95
? FILING DATE: 28-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Miller, Charles E.
? REGISTRATION NUMBER: 24,576
? REFERENCE/DOCKET NUMBER: 7426-043-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 488 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? OS-08-809-326A-1

```

Query Match	Score	DB 4	Length
98.28;	487;	DB 4;	Length 488

4 mcaaaacccgdcnagmmsqvelstipqgvpqddkscgnetnqiioqtrgskntemesdatiag 63

1 MSISSSSGPDNKNIMSQVLTSIPQGVPPQDDKLSGNETKQIQOTRQGNKNTMESDATING 60

64 ASGKDKTSSTKTETAPQQGVAAKESSESQKAGADTGVSGAAATTASNTATKIAMQTSI 123

D6 61 ASGDKTSTTKETAPQGVAAKGESSESQKAGADTGVSGAATTASNTATKIAMOTSI 120

QY 124 EEASKSMESTLESIOSI.SAOMKEVEAVVVAI CCKSCCSAVT EMBET BYDQWEMDQWU- 100

[illegible]

121 LEASNSMESILESLSAQMKEEAVVAALSCKSSSAKLETPPELPPGVTPRSEVI 180

QY 184 EIGLALAKAIQTIGEATKSALSNYASTQADQTNKLGLEKQAIKIDKEREYQEMKAAE 243

Db 181 EIGLALAKAIQTIGEATKSALSASTQAQADQTNKLGLEKQAIKIDKEREYOEKAAE 240

QY 244 QKSKLEGTMDTVNTVMIAVSAITVISTVAATETCGAGTACIAGAATWCACCCACCA 303

341  
b6  
b7C  
b7D  
b7E  
b7F  
b7G  
b7H  
b7I  
b7J  
b7K  
b7L  
b7M  
b7N  
b7O  
b7P  
b7Q  
b7R  
b7S  
b7T  
b7U  
b7V  
b7W  
b7X  
b7Y  
b7Z  
b7AA  
b7AB  
b7AC  
b7AD  
b7AE  
b7AF  
b7AG  
b7AH  
b7AI  
b7AJ  
b7AK  
b7AL  
b7AM  
b7AN  
b7AO  
b7AP  
b7AQ  
b7AR  
b7AS  
b7AT  
b7AU  
b7AV  
b7AW  
b7AX  
b7AY  
b7AZ  
b7BA  
b7BB  
b7BC  
b7BD  
b7BE  
b7BF  
b7BG  
b7BH  
b7BI  
b7BJ  
b7BK  
b7BL  
b7BM  
b7BN  
b7BO  
b7BP  
b7BQ  
b7BR  
b7BS  
b7BT  
b7BU  
b7BV  
b7BW  
b7BX  
b7BY  
b7BZ  
b7CA  
b7CB  
b7CC  
b7CD  
b7CE  
b7CF  
b7CG  
b7CH  
b7CI  
b7CJ  
b7CK  
b7CL  
b7CM  
b7CN  
b7CO  
b7CP  
b7CQ  
b7CR  
b7CS  
b7CT  
b7CU  
b7CV  
b7CW  
b7CX  
b7CY  
b7CZ  
b7DA  
b7DB  
b7DC  
b7DD  
b7DE  
b7DF  
b7DG  
b7DH  
b7DI  
b7DJ  
b7DK  
b7DL  
b7DM  
b7DN  
b7DO  
b7DP  
b7DQ  
b7DR  
b7DS  
b7DT  
b7DU  
b7DV  
b7DW  
b7DX  
b7DY  
b7DZ  
b7EA  
b7EB  
b7EC  
b7ED  
b7EE  
b7EF  
b7EG  
b7EH  
b7EI  
b7EJ  
b7EK  
b7EL  
b7EM  
b7EN  
b7EO  
b7EP  
b7EQ  
b7ER  
b7ES  
b7ET  
b7EU  
b7EV  
b7EW  
b7EX  
b7EY  
b7EZ  
b7FA  
b7FB  
b7FC  
b7FD  
b7FE  
b7FF  
b7FG  
b7FH  
b7FI  
b7FJ  
b7FK  
b7FL  
b7FM  
b7FN  
b7FO  
b7FP  
b7FQ  
b7FR  
b7FS  
b7FT  
b7FU  
b7FV  
b7FW  
b7FX  
b7FY  
b7FZ  
b7GA  
b7GB  
b7GC  
b7GD  
b7GE  
b7GF  
b7GG  
b7GH  
b7GI  
b7GJ  
b7GK  
b7GL  
b7GM  
b7GN  
b7GO  
b7GP  
b7GQ  
b7GR  
b7GS  
b7GT  
b7GU  
b7GV  
b7GW  
b7GX  
b7GY  
b7GZ  
b7HA  
b7HB  
b7HC  
b7HD  
b7HE  
b7HF  
b7HG  
b7HH  
b7HI  
b7HJ  
b7HK  
b7HL  
b7HM  
b7HN  
b7HO  
b7HP  
b7HQ  
b7HR  
b7HS  
b7HT  
b7HU  
b7HV  
b7HW  
b7HX  
b7HY  
b7HZ  
b7IA  
b7IB  
b7IC  
b7ID  
b7IE  
b7IF  
b7IG  
b7IH  
b7IJ  
b7IK  
b7IL  
b7IM  
b7IN  
b7IO  
b7IP  
b7IQ  
b7IR  
b7IS  
b7IT  
b7IU  
b7IV  
b7IW  
b7IX  
b7IY  
b7IZ  
b7JA  
b7JB  
b7JC  
b7JD  
b7JE  
b7JF  
b7JG  
b7JH  
b7JI  
b7JJ  
b7JK  
b7JL  
b7JM  
b7JN  
b7JO  
b7JP  
b7JQ  
b7JR  
b7JS  
b7JT  
b7JU  
b7JV  
b7JW  
b7JX  
b7JY  
b7JZ  
b7KA  
b7KB  
b7KC  
b7KD  
b7KE  
b7KF  
b7KG  
b7KH  
b7KI  
b7KJ  
b7KK  
b7KL  
b7KM  
b7KN  
b7KO  
b7KP  
b7KQ  
b7KR  
b7KS  
b7KT  
b7KU  
b7KV  
b7KW  
b7KX  
b7KY  
b7KZ  
b7LA  
b7LB  
b7LC  
b7LD  
b7LE  
b7LF  
b7LG  
b7LH  
b7LI  
b7LJ  
b7LK  
b7LL  
b7LM  
b7LN  
b7LO  
b7LP  
b7LQ  
b7LR  
b7LS  
b7LT  
b7LU  
b7LV  
b7LW  
b7LX  
b7LY  
b7LZ  
b7MA  
b7MB  
b7MC  
b7MD  
b7ME  
b7MF  
b7MG  
b7MH  
b7MI  
b7MJ  
b7MK  
b7ML  
b7MN  
b7MO  
b7MP  
b7MQ  
b7MR  
b7MS  
b7MT  
b7MU  
b7MV  
b7MW  
b7MX  
b7MY  
b7MZ  
b7NA  
b7NB  
b7NC  
b7ND  
b7NE  
b7NF  
b7NG  
b7NH  
b7NI  
b7NJ  
b7NK  
b7NL  
b7NM  
b7NO  
b7NP  
b7NQ  
b7NR  
b7NS  
b7NT  
b7NU  
b7NV  
b7NW  
b7NX  
b7NY  
b7NZ  
b7OA  
b7OB  
b7OC  
b7OD  
b7OE  
b7OF  
b7OG  
b7OH  
b7OI  
b7OJ  
b7OK  
b7OL  
b7OM  
b7ON  
b7OO  
b7OP  
b7OQ  
b7OR  
b7OS  
b7OT  
b7OU  
b7OV  
b7OW  
b7OX  
b7OY  
b7OZ  
b7PA  
b7PB  
b7PC  
b7PD  
b7PE  
b7PF  
b7PG  
b7PH  
b7PI  
b7PJ  
b7PK  
b7PL  
b7PM  
b7PN  
b7PO  
b7PP  
b7PQ  
b7PR  
b7PS  
b7PT  
b7PU  
b7PV  
b7PW  
b7PX  
b7PY  
b7PZ  
b7QA  
b7QB  
b7QC  
b7QD  
b7QE  
b7QF  
b7QG  
b7QH  
b7QI  
b7QJ  
b7QK  
b7QL  
b7QM  
b7QN  
b7QO  
b7QP  
b7QQ  
b7QR  
b7QS  
b7QT  
b7QU  
b7QV  
b7QW  
b7QX  
b7QY  
b7QZ  
b7RA  
b7RB  
b7RC  
b7RD  
b7RE  
b7RF  
b7RG  
b7RH  
b7RI  
b7RJ  
b7RK  
b7RL  
b7RM  
b7RN  
b7RO  
b7RP  
b7RQ  
b7RR  
b7RS  
b7RT  
b7RU  
b7RV  
b7RW  
b7RX  
b7RY  
b7RZ  
b7SA  
b7SB  
b7SC  
b7SD  
b7SE  
b7SF  
b7SG  
b7SH  
b7SI  
b7SJ  
b7SK  
b7SL  
b7SM  
b7SN  
b7SO  
b7SP  
b7SQ  
b7SR  
b7SS  
b7ST  
b7SU  
b7SV  
b7SW  
b7SX  
b7SY  
b7SZ  
b7TA  
b7TB  
b7TC  
b7TD  
b7TE  
b7TF  
b7TG  
b7TH  
b7TI  
b7TJ  
b7TK  
b7TL  
b7TM  
b7TN  
b7TO  
b7TP  
b7TQ  
b7TR  
b7TS  
b7TT  
b7TU  
b7TV  
b7TW  
b7TX  
b7TY  
b7TZ  
b7UA  
b7UB  
b7UC  
b7UD  
b7UE  
b7UF  
b7UG  
b7UH  
b7UI  
b7UJ  
b7UK  
b7UL  
b7UM  
b7UN  
b7UO  
b7UP  
b7UQ  
b7UR  
b7US  
b7UT  
b7UU  
b7UV  
b7UW  
b7UX  
b7UY  
b7UZ  
b7VA  
b7VB  
b7VC  
b7VD  
b7VE  
b7VF  
b7VG  
b7VH  
b7VI  
b7VJ  
b7VK  
b7VL  
b7VM  
b7VN  
b7VO  
b7VP  
b7VQ  
b7VR  
b7VS  
b7VT  
b7VU  
b7VV  
b7VW  
b7VX  
b7VY  
b7VZ  
b7WA  
b7WB  
b7WC  
b7WD  
b7WE  
b7WF  
b7WG  
b7WH  
b7WI  
b7WJ  
b7WK  
b7WL  
b7WM  
b7WN  
b7WO  
b7WP  
b7WQ  
b7WR  
b7WS  
b7WT  
b7WU  
b7WV  
b7WW  
b7WX  
b7WY  
b7WZ  
b7XA  
b7XB  
b7XC  
b7XD  
b7XE  
b7XF  
b7XG  
b7XH  
b7XI  
b7XJ  
b7XK  
b7XL  
b7XM  
b7XN  
b7XO  
b7XP  
b7XQ  
b7XR  
b7XS  
b7XT  
b7XU  
b7XV  
b7XW  
b7XX  
b7XY  
b7XZ  
b7YA  
b7YB  
b7YC  
b7YD  
b7YE  
b7YF  
b7YG  
b7YH  
b7YI  
b7YJ  
b7YK  
b7YL  
b7YM  
b7YN  
b7YO  
b7YP  
b7YQ  
b7YR  
b7YS  
b7YT  
b7YU  
b7YV  
b7YW  
b7YX  
b7YY  
b7YZ  
b7ZA  
b7ZB  
b7ZC  
b7ZD  
b7ZE  
b7ZF  
b7ZG  
b7ZH  
b7ZI  
b7ZJ  
b7ZK  
b7ZL  
b7ZM  
b7ZN  
b7ZO  
b7ZP  
b7ZQ  
b7ZR  
b7ZS  
b7ZT  
b7ZU  
b7ZV  
b7ZW  
b7ZX  
b7ZY  
b7ZZ  
b7AA  
b7AB  
b7AC  
b7AD  
b7AE  
b7AF  
b7AG  
b7AH  
b7AI  
b7AJ  
b7AK  
b7AL  
b7AM  
b7AN  
b7AO  
b7AP  
b7AQ  
b7AR  
b7AS  
b7AT  
b7AU  
b7AV  
b7AW  
b7AX  
b7AY  
b7AZ  
b7BA  
b7BB  
b7BC  
b7BD  
b7BE  
b7BF  
b7BG  
b7BH  
b7BI  
b7BJ  
b7BK  
b7BL  
b7BM  
b7BN  
b7BO  
b7BP  
b7BQ  
b7BR  
b7BS  
b7BT  
b7BU  
b7BV  
b7BW  
b7BX  
b7BY  
b7BZ  
b7CA  
b7CB  
b7CC  
b7CD  
b7CE  
b7CF  
b7CG  
b7CH  
b7CI  
b7CJ  
b7CK  
b7CL  
b7CM  
b7CN  
b7CO  
b7CP  
b7CQ  
b7CR  
b7CS  
b7CT  
b7CU  
b7CV  
b7CW  
b7CX  
b7CY  
b7CZ  
b7DA  
b7DB  
b7DC  
b7DD  
b7DE  
b7DF  
b7DG  
b7DH  
b7DI  
b7DJ  
b7DK  
b7DL  
b7DM  
b7DN  
b7DO  
b7DP  
b7DQ  
b7DR

241 QKSNDEGIMDI VNI VMIAVSVAILVISI VAAIFTCGAGLAGLAAGAAVGAAGGAAGA 300

304 A A A T T V A T Q I T V Q A V V Q A V K Q A V I T A V R Q A I T A A I K A V K S G I K A F I K T L V K A I A K A I S K 363

Db 301 AAATTVATQITVQAVVQAVKQAVITAVRQAITAATKAAVKSGIKAFIKTLVKAIAKAISK 360

364 GISKVFAKGTOMIAKNFKLSKVISSITSKWVTGVGVVAPACGKTMOMOTSEUCOY 433

[illegible]

01 015NVEPNOIEMIANNEPRLSKVISSLSKMWLVGVGVVAAAPALGKIGIMQMQLSEMQON 420

424 VAQFOKEVGKLOAADMISMFTQFWQASKIASKOTGESNEMTOKATKLGACILKAYAAI 483

db 421 VAQFQKEVGKLAADMI SMFTQFWQASKIASKQTGESNEMTOKATKLGAI LKAYAI 480

484 SGATAGA 490

$\frac{1}{2}$

401 BGALAGA 48 /

## RESULT 2

JS-08-809-326A-15

sequence 15, Application US/08809326A  
Patent No. 6165478

GENERAL INFORMATION:

APPLICANT: Izutsu, Hiroshi  
APPLICANT: Obara, Kazuhiko



APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR  
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106011/95  
FILING DATE: 28-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Charles E.  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 7426-043-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEO ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 649 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-809-326A-15

Query Match 98.28; Score 487; DB 4; Length 649;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSSGPNOKIMSOVLTSTPOGVPOODKLSGNETKOTIOROGKNTMESDATTAG 63  
DB 162 MSISSSSGPNOKIMSOVLTSTPOGVPOODKLSGNETKOTIOROGKNTMESDATTAG 221  
QY 64 ASGRDKTSSTTKETAPAOOGVAACKESSESOKAGADTGVSGAAATTSNTATKIAMQTSI 123  
DB 222 ASGRDKTSSTTKETAPAOOGVAACKESSESOKAGADTGVSGAAATTSNTATKIAMQTSI 281

QY 124 EASKSMESTLESLOSISAQMKVEAVVAALSGKSGSAKLEPPEPKRGVPRSPVI 183  
DB 282 EASKSMESTLESLOSISAQMKVEAVVAALSGKSGSAKLEPPEPKRGVPRSPVI 341  
QY 184 EIGALAKAIOTLEATRSALSNASTQADQTNKLGLEQAIKIDEREYDEMKAEE 243  
DB 342 EIGALAKAIOTLEATRSALSNASTQADQTNKLGLEQAIKIDEREYDEMKAEE 401  
QY 244 OKSDLEGTMDTVNTVIAVSVAITVISVAIIFTCGAGLAGLAAGAAVAAAAGAGA 303  
DB 402 OKSDLEGTMDTVNTVIAVSVAITVISVAIIFTCGAGLAGLAAGAAVAAAAGAGA 461  
QY 304 AAATTVATQTTVOAVQAVKAVITVAVQAITTAIKAIVKSGIAFTLVKATAIKAS 363  
DB 462 AAATTVATQTTVOAVQAVKAVITVAVQAITTAIKAIVKSGIAFTLVKATAIKAS 521  
QY 364 GISKVFAGTOMIAKNPFLSKVISLSLTSKWVYGVVVAAPALGKIMOMOLEMOON 423  
DB 522 GISKVFAGTOMIAKNPFLSKVISLSLTSKWVYGVVVAAPALGKIMOMOLEMOON 581  
QY 424 VAOFQEVGKLOAAADMISMTQFMOQASKIASKOTGESNEMTQKATLGAQILKAYAI 483  
DB 582 VAOFQEVGKLOAAADMISMTQFMOQASKIASKOTGESNEMTQKATLGAQILKAYAI 641  
QY 484 SGATAGA 490  
DB 642 SGATAGA 648

RESULT 3  
US-08-809-326A-5  
Sequence 5, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Izutsu, Kazuhiro  
APPLICANT: Obara, Kazuhiko  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95

```

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-809-326A-2

Query Match 52.2%; Score 259; DB 4; Length 271;
Best Local Similarity 100.0%; Freq. No. 1,6e-226;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

US-08-809-326A-16  
 ; Sequence 16, Application US/08809326A  
 ; Patent No. 6165478  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Izutsu, Hiroshi  
 ; APPLICANT: Obara, Kazuhiko  
 ; APPLICANT: Matsumoto, Akira  
 ; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
 ; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 ; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS  
 ; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
 ; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
 ; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
 ; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR  
 ; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
 ; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/809,326A  
 ; FILING DATE: 19-MAR-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 224711/94  
 ; FILING DATE: 20-SEP-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 106006/95  
 ; FILING DATE: 28-APR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 106008/95  
 ; FILING DATE: 28-APR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 106009/95  
 ; FILING DATE: 28-APR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 106010/95  
 ; FILING DATE: 28-APR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 106011/95  
 ; FILING DATE: 28-APR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Miller, Charles E.  
 ; REGISTRATION NUMBER: 24,576  
 ; REFERENCE/DOCKET NUMBER: 7426-043-999  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-8864/9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 432 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-809-326A-16

Query Match 52.2%; Score 259; DB 4; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-226;  
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 MSISSSGPDNKNIMOVLTSTPGVPOQDKLSGNETKIOQTRQGNTEMESDATTAG 63

Db 162 MSISSSSGPDNKNIMOVLTSTPGVPOQDKLSGNETKIOQTRQGNTEMESDATTAG 221  
 Oy 64 ASGDKDTSSSTTKTETAPQOQVAAKESSESQKAGADTCVGAATVTSNTATKIAMQTSI 123  
 Db 222 ASGDKDTSSSTTKTETAPQOQVAAKESSESQKAGADTCVGAATVTSNTATKIAMQTSI 281  
 Oy 124 EEAQSMSTLESLSAQAOMKEVAVVAALSGSSGSAKLETPPKPGVTPRSFVI 183  
 Db 282 EEAQSMSTLESLSAQAOMKEVAVVAALSGSSGSAKLETPPKPGVTPRSFVI 341  
 Oy 184 EIGLAKAIDTLEBATSALSNASTOAOADOTNKILGLEQATKIDKEREYQEMKAAE 243  
 Db 342 EIGLAKAIDTLEBATSALSNASTOAOADOTNKILGLEQATKIDKEREYQEMKAAE 401  
 Oy 244 QKSKDLEGTMDTVNTVMIA 262  
 Db 402 QKSKDLEGTMDTVNTVMIA 420

RESULT 6  
 ; US-09-135-994-11  
 ; Sequence 11, Application US/09135994A  
 ; Patent No. 6280938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rannum et al.  
 ; TITLE OF INVENTION: SCAT GENE AND METHODS OF USE  
 ; FILE REFERENCE: University of Minnesota  
 ; CURRENT APPLICATION NUMBER: US/09/135,994A  
 ; CURRENT FILING DATE: 1998-08-18  
 ; EARLIER APPLICATION NUMBER: 60/056,170  
 ; EARLIER FILING DATE: 1997-08-19  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 11  
 ; LENGTH: 27  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-135-994-11

Query Match 1.6%; Score 8; DB 4; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 294 AAAGGAA 301  
 Db 17 AAAGGAA 24

RESULT 7  
 ; US-08-425-069-50  
 ; Sequence 50, Application US/08425069  
 ; Patent No. 5728810  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lewis, Randolph V.  
 ; APPLICANT: Xu, Ming  
 ; APPLICANT: Himan, Michael B.  
 ; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
 ; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
 ; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
 ; NUMBER OF SEQUENCES: 69  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
 ; STREET: 301 No. 5728810th Washington Street  
 ; CITY: Falls Church  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22046  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/425,069
  FILING DATE: 19-APR-1995
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Murphy Jr., Gerald M
    REGISTRATION NUMBER: 28,977
    REFERENCE/DOCKET NUMBER: 1447-106P
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703) 205-8000
      TELEFAX: (703) 205-8050
  TELEX:
  INFORMATION FOR SEQ ID NO: 50:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 28 amino acids
      TYPE: amino acid
      STRANDEDNESS: not relevant
      TOPOLOGY: not relevant
    MOLECULE TYPE: peptide
    HYPOTHETICAL: NO
    FRAGMENT TYPE: Internal
    ORIGINAL SOURCE:
      ORGANISM: nephila clavipes
  FEATURE:
    NAME/KEY: Peptide
    LOCATION: 1..28
    OTHER INFORMATION: /label= silk1_repeat
US-08-425-069-50

Query Match      1.6%; Score 8; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
   |||||
Db 12 GAAAGGA 19

RESULT 8
US-08-317-844B-50
; Sequence 50, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
```

```

TELEX: 248345
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: nephila clavipes
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..28
; OTHER INFORMATION: /label= silk1_repeat
US-08-317-844B-50

Query Match      1.6%; Score 8; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
   |||||
Db 12 GAAAGGA 19

RESULT 9
US-08-425-069-61
; Sequence 61, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,069
; FILING DATE: 19-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
```

ORGANISM: nephila clavipes  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..30  
OTHER INFORMATION: /label= silk1\_repeat  
US-08-425-069-61

Query Match 1.6%; Score 8; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300  
|||||||  
DB 12 GAAAGCA 19

RESULT 10  
US-08-317-844B-61  
Sequence 61, Application US/08317844B  
Patent No. 5989894  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5989894th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22046

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/317,844B  
FILING DATE: 04-OCT-1994  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345

INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: nephila clavipes  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..30  
OTHER INFORMATION: /label= silk1\_repeat  
US-08-317-844B-61

Query Match 1.6%; Score 8; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300  
|||||||  
DB 12 GAAAGCA 19

RESULT 11  
US-08-425-069-6  
Sequence 6, Application US/08425069  
Patent No. 5728810  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5728810th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22046

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,069  
FILING DATE: 19-APR-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-106P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX:

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Nephila clavipes  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..33  
OTHER INFORMATION: /label= representative  
OTHER INFORMATION: /note="this peptide is a representative one that  
OTHER INFORMATION: illustrates the g9x9yg hexamer repeat motif of the  
OTHER INFORMATION: spider silk protein I."

US-08-425-069-6  
Query Match 1.6%; Score 8; DB 1; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300  
|||||||  
DB 6 GAAAGCA 13

RESULT 12  
US-08-317-844B-6  
Sequence 6, Application US/08317844B  
Patent No. 5989894  
GENERAL INFORMATION:

```

; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Nephila clavipes
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..33
; OTHER INFORMATION: /label=representative
; OTHER INFORMATION: /note="this peptide is a representative one that
; OTHER INFORMATION: illustrates the 99x9y9 hexamer repeat motif of the
; OTHER INFORMATION: spider silk protein I."
; US-08-317-844B-6
;
; Query Match 1.6%; Score 8; DB 2; Length 33;
; Best Local Similarity 100.0%; Pred. No. 3.2;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 293 GAAAGGA 300
; DB 6 GAAAGGA 13
;
; RESULT 13
; US-08-185-414E-2
; Sequence 2, Application US/08185414E
; Patent No. 5556953
; GENERAL INFORMATION:
; APPLICANT: Zhang, Lei
; APPLICANT: Vijay, Hari M.
; APPLICANT: Rode, Harold
; TITLE OF INVENTION: ALLERGEN OF CLADOSPORIUM HERBARUM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: George A. Seaby
; ADDRESS: Seaby and Maclean
; STREET: 880 Wellington Street, Suite 708
; CITY: Ottawa

```

```

; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,414E
; FILING DATE: January 24, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George A. Seaby
; REGISTRATION NUMBER: 24,034
; REFERENCE/DOCKET NUMBER: 1747
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 232-5815
; TELEFAX: (613) 232-5831
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 111 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Ch2.1 protein
; LOCATION: 1..111
; US-08-185-414E-2

```

```

; Query Match 1.6%; Score 8; DB 1; Length 111;
; Best Local Similarity 100.0%; Pred. No. 11;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 293 GAAAGGA 300
; DB 78 GAAAGGA 85

```

```

; RESULT 14
; US-09-135-994-12
; Sequence 12, Application US/09135994A
; Patent No. 6280938
; GENERAL INFORMATION:
; APPLICANT: Rannum et al.
; TITLE OF INVENTION: SCAT GENE AND METHODS OF USE
; FILE REFERENCE: University of Minnesota
; CURRENT APPLICATION NUMBER: US/09/135,994A
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,170
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 12
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-135-994-12

```

```

; Query Match 1.6%; Score 8; DB 4; Length 129;
; Best Local Similarity 100.0%; Pred. No. 12;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 294 AAAAGGA 301
; DB 17 AAAAGGA 24

```

```

; RESULT 15
; US-08-911-319A-1
; Sequence 1, Application US/08911319A
; Patent No. 5968798
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.

```

APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/911,319A  
FILING DATE: August 14, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Muenzen, Colette C.  
REGISTRATION NUMBER: 39,784  
REFERENCE/DOCKET NUMBER: PF-0363 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 164 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THPINOT03  
CLONE: 2447829  
US-08-911-319A-1

Query Match 1.6%; Score 8; DB 2; Length 164;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAAGAAA 306  
Db 29 GAAGAAA 36

RESULT 16  
US-09-352-619-1  
Sequence 1, Application US/09352619  
Patent No. 6084070  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/352,619

FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/911,319  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Muenzen, Colette C.  
REGISTRATION NUMBER: 39,784  
REFERENCE/DOCKET NUMBER: PF-0363 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 164 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THPINOT03  
CLONE: 2447829  
US-09-352-619-1

Query Match 1.6%; Score 8; DB 3; Length 164;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAAGAAA 306  
Db 29 GAAGAAA 36

RESULT 17  
US-08-775-009-37  
Sequence 37, Application US/08775009  
Patent No. 5935783  
GENERAL INFORMATION:  
APPLICANT: Gong, Wellong  
APPLICANT: Emanuel, Beverly S.  
APPLICANT: Budarf, Marcia L.  
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and  
TITLE OF INVENTION: Velocardiofacial Syndrome Minimal Critical Region  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock washburn Kurtz Mackiewicz &  
ADDRESSEE: No. 5935783rls, LLP  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/775,009  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Trujillo, Doreen Yatkoo  
REGISTRATION NUMBER: 35,719  
REFERENCE/DOCKET NUMBER: CH-0681  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3439  
TELEFAX: (215) 568-3100  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 205 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-775-009-37

## Query Match

1.6%; Score 8; DB 2; Length 205;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGGA 301  
|||||||

Db 2 AAAAGGA 9

## RESULT 18

US-08-364-081-3  
Sequence 3, Application US/08364081  
Patent No. 5585464

## GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, No. 5585464th &amp; Western

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: AST Advantage NB-SX20

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/364,081

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,780

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553

REFERENCE/DOCKET NUMBER: T781CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (801)566-6633

TELEFAX: (801)566-0750

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-364-081-3

## Query Match

1.6%; Score 8; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
|||||||

Db 260 GAAAGGA 267

## RESULT 19

US-08-630-552-3  
Sequence 3, Application US/08630552  
Patent No. 572314

## GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, No. 5723314th &amp; Western

STREET: 9035 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: AST Advantage NB-SX20

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,552

FILING DATE: 10-APR-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/364,081

FILING DATE:

APPLICATION NUMBER: 08/019,780

FILING DATE: 19-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553

REFERENCE/DOCKET NUMBER: T781CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (801)566-6633

TELEFAX: (801)566-0750

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-630-552-3

## Query Match

1.6%; Score 8; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
|||||||

Db 260 GAAAGGA 267

## RESULT 20

PCT-US95-16558-3  
Sequence 3, Application PC/TUS9516558  
GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, North &amp; Western

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb

COMPUTER: IBM ThinkPad 340

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16558

FILING DATE: 27-DEC-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,780

FILING DATE: 19-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553



REFERENCE/DOCKET NUMBER: T781CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (801)566-6633  
TELEFAX: (801)566-0750  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 331 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-16558-3

Query Match 1.6%; Score 8; DB 5; Length 331;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
Db 260 GAAAGGA 267

RESULT 21  
US-09-347-801-18  
Sequence 18, Application US/09347801  
Patent No. 6262345  
GENERAL INFORMATION:  
APPLICANT: Allen, Steve  
APPLICANT: Lee, Jian Ming  
TITLE OF INVENTION: Plant Protein Kinases  
FILE REFERENCE: BB-1171  
CURRENT APPLICATION NUMBER: US/09/347,801  
CURRENT FILING DATE: 1999-07-02  
EARLIER APPLICATION NUMBER: 60/092,438  
EARLIER FILING DATE: July 10, 1998  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 18  
LENGTH: 625  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-347-801-18

Query Match 1.6%; Score 8; DB 4; Length 625;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 AGAAGGA 295  
Db 144 AGAAGGA 151

RESULT 22  
US-08-556-978B-19  
Sequence 19, Application US/08556978B  
Patent No. 6268169  
GENERAL INFORMATION:  
APPLICANT: FAHNESTOCK, STEPHEN F.  
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
NUMBER OF INVENTION: SPIDER SILK ANALOGS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WINDOWS 95  
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/556,978B

FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/077,600  
FILING DATE: JUNE 15, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CR-9389-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 651 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-556-978B-19

Query Match 1.6%; Score 8; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
Db 317 GAAAGGA 324

RESULT 23  
US-09-247-806-1  
Sequence 1, Application US/09247806  
Patent No. 6280747  
GENERAL INFORMATION:  
APPLICANT: PHILIPPE, Michel  
APPLICANT: GARSON, Jean-Claude  
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT  
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN  
FILE REFERENCE: 6388-0365-0  
CURRENT APPLICATION NUMBER: US/09/247,806  
CURRENT FILING DATE: 1999-02-11  
EARLIER APPLICATION NUMBER: FR 98/01614  
EARLIER FILING DATE: 1998-02-11  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 651  
TYPE: PRT  
ORGANISM: Nephila clavipes  
US-09-247-806-1

Query Match 1.6%; Score 8; DB 4; Length 651;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
Db 317 GAAAGGA 324

RESULT 24  
US-08-425-069-2  
Sequence 2, Application US/08425069  
Patent No. 5728810  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
APPLICANT: Himan, Michael B.  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL

;; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
;; STREET: 301 No. 5728010th Washington Street  
;; CITY: Falls Church  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22046  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/425,069  
;; FILING DATE: 19-APR-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Murphy Jr., Gerald M  
;; REGISTRATION NUMBER: 28,977  
;; REFERENCE/DOCKET NUMBER: 1447-106P  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 205-8000  
;; TELEFAX: (703) 205-8050  
;; TELEX:  
;;  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 718 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-425-069-2

Query Match  
Best Local Similarity 1.6%; Score 8; DB 1; Length 718;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
|||  
Db 317 GAAAGGA 324

RESULT 25  
US-08-317-844B-2  
; Sequence 2, Application US/08317844B  
; Patent No. 5989894  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, Randolph V.  
; APPLICANT: Xu, Ming  
; APPLICANT: Himan, Michael B.  
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 No. 5989894th Washington Street  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22046  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/317,844B  
; FILING DATE: 04-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M

;; REGISTRATION NUMBER: 28,977  
;; REFERENCE/DOCKET NUMBER: 1447-105P  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 241-1300  
;; TELEFAX: (703) 241-2848  
;; TELEX: 248345  
;;  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 718 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-317-844B-2

Query Match  
Best Local Similarity 1.6%; Score 8; DB 2; Length 718;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300  
|||  
Db 317 GAAAGGA 324

RESULT 26  
US-09-034-177-3  
; Sequence 3, Application US/09034177  
; Patent No. 6127146  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN FIBROUS PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,177  
; FILING DATE: HEREMITH  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0486 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 747 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: GI 1174414  
; US-09-034-177-3

Query Match  
Best Local Similarity 1.6%; Score 8; DB 3; Length 747;  
Matches 100.0%; Pred. No. 69;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300  
|1111111|  
Db 624 GAAAGCA 631

RESULT 27  
US-08-630-915A-37

; Sequence 37, Application US/08630915A  
; Patent No. 6309820

; GENERAL INFORMATION:

; APPLICANT: SPARKS, Andrew B.

; APPLICANT: HOFEMAN, No. 6309820h

; APPLICANT: KAY, Brian K.

; APPLICANT: FOWLES, Dana M.

; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL

; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND

; NUMBER OF SEQUENCES: 227

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; City: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/630,915A

; FILING DATE: 03-APR-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-174

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-8864/9741

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1400 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; US-08-630-915A-37

; Query Match

; Best Local Similarity 100.0%; Pred. No. 1.3e+02;

; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGAA 301  
|1111111|  
Db 307 AAAAGAA 314

RESULT 28  
US-09-548-372D-13

; Sequence 13, Application US/09548372D

; Patent No. 6420534

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/62801

; CURRENT APPLICATION NUMBER: US/09/548,372D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 13

; LENGTH: 2088

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-548-372D-13

; Query Match

; Best Local Similarity 100.0%; Pred. No. 1.9e+02;

; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 297 AGAAGAA 304  
|1111111|  
Db 692 AGAAGAA 699

RESULT 29  
US-09-548-367D-13

; Sequence 13, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 13

; LENGTH: 2088

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-548-367D-13

; Query Match

; Best Local Similarity 100.0%; Pred. No. 1.9e+02;

; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 297 AGAAGAA 304  
|1111111|  
Db 692 AGAAGAA 699

RESULT 30  
US-09-738-884-1

; Sequence 1, Application US/09738884

; Patent No. 6391606

; GENERAL INFORMATION:

; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: C1000849

; CURRENT APPLICATION NUMBER: US/09/738,884

```
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2211
; TYPE: PRT
; ORGANISM: Human
US-09-738-884-1
```

```
Query Match          1.6%; Score 8; DB 4; Length 2211;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 297 AGGAGAA 304
Db 1361 AGGAGAA 1368
```

```
RESULT 31
US-09-627-650B-7
; Sequence 7, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-7
```

```
Query Match          1.6%; Score 8; DB 4; Length 2508;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 32
US-09-436-063C-7
; Sequence 7, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-7
```

```
Query Match          1.6%; Score 8; DB 4; Length 2508;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 33
US-09-627-650B-3
; Sequence 3, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-3
```

```
Query Match          1.6%; Score 8; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 34
US-09-436-063C-3
; Sequence 3, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-3
```

```
Query Match          1.6%; Score 8; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 35
US-09-627-650B-9
; Sequence 9, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-9

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 2601;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191

RESULT 36
US-09-436-063C-9
; Sequence 9, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-9

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 2601;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191

RESULT 37
US-09-336-447A-76
; Sequence 76, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
```

```
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 76
; LENGTH: 3788
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (1036)..(3786)
; OTHER INFORMATION: Xaa = any
US-09-336-447A-76

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 3788;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 3603 GAAGAAA 3610

RESULT 38
5178861-16
; Patent No. 5178861
; APPLICANT: VERGARA, ULISES;RUIZ, ANDRES;FERREIRA, ARTURO;
; NUSSENZWEIG, RUTH S.;NUSSENZWEIG, VICTOR N.
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES
; OF CIRCUMPOROZOITE PROTEINS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,241
; FILING DATE: 22-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO:16
; LENGTH: 11
5178861-16

Query Match
Best Local Similarity 1.4%; Score 7; DB 6; Length 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGC 299
Db 2 GAAAGC 8

RESULT 39
US-08-425-069-5
; Sequence 5, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Himan, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia
```

COUNTRY: U.S.A.  
ZIP: 22046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,069  
FILING DATE: 19-APR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-106P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Nephila clavipes  
FEATURE:  
NAME/KEY: Duplication  
LOCATION: 1..21  
OTHER INFORMATION: /label=repeat\_unit  
OTHER INFORMATION: /note="spider silk protein repeat unit"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /label=alanine\_stretch  
OTHER INFORMATION: /note="this segment of alanines in the repeat  
OTHER INFORMATION: unit can also contain 7 alanine residues."  
US-08-425-069-5  
Query Match 1.4%; Score 7; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 294 AAAAGGA 300  
DB 3 AAAAGGA 9  
RESULT 40  
US-08-317-844B-5  
Sequence 5, Application US/08317844B  
Patent No. 5989894  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
APPLICANT: Hinman, Michael B.  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5989894th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/317,844B  
FILING DATE: 04-OCT-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Nephila clavipes  
FEATURE:  
NAME/KEY: Duplication  
LOCATION: 1..21  
OTHER INFORMATION: /label=repeat\_unit  
OTHER INFORMATION: /note="spider silk protein repeat unit"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..6  
OTHER INFORMATION: /label=alanine\_stretch  
OTHER INFORMATION: /note="this segment of alanines in the repeat  
OTHER INFORMATION: unit can also contain 7 alanine residues."  
US-08-317-844B-5  
Query Match 1.4%; Score 7; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 294 AAAAGGA 300  
DB 3 AAAAGGA 9  
RESULT 41  
US-07-987-286-3  
Sequence 3, Application 07/987286  
Patent No. 5500366  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, GREGORY J  
APPLICANT: GECZY, ANDREW F  
TITLE OF INVENTION: T-CELL EPITOPES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 07/987,286  
FILING DATE: 16-MAR-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361

FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-987-286-3

Query Match 1.4%; Score 7; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAA 295  
Db 10 GAAVGAA 16

RESULT 42  
US-08-614-626-3  
Sequence 3, Application US/08614626  
Patent No. 5928644  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, GREGORY J  
APPLICANT: GECZY, ANDREW F  
TITLE OF INVENTION: T-CELL EPITOPES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/614,626  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/987,286  
FILING DATE: 16-MAR-1993  
APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361  
FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-614-626-3

Query Match 1.4%; Score 7; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAA 295  
Db 10 GAAVGAA 16

RESULT 43  
US-07-987-286-11  
Sequence 11, Application 07/987286  
Patent No. 5500366  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, GREGORY J  
APPLICANT: GECZY, ANDREW F  
TITLE OF INVENTION: T-CELL EPITOPES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 07/987,286  
FILING DATE: 16-MAR-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361  
FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note= "RESIDUE 1 IS PYROGLUTAMIC  
OTHER INFORMATION: ACID"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 25  
OTHER INFORMATION: /note= "RESIDUE 25 IS CYSTEINAMIDE"  
US-07-987-286-11

Query Match 1.4%; Score 7; DB 1; Length 25;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAA 295  
Db 11 GAAVGAA 17

RESULT 44  
US-07-987-286-24  
Sequence 24, Application 07/987286  
Patent No. 5500566  
GENERAL INFORMATION:  
APPLICANT: RUSSELL-JONES, GREGORY J  
APPLICANT: GECZY, ANDREW F  
TITLE OF INVENTION: T-CELL EPTIOPES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 07/987,286  
FILING DATE: 16-MAR-1993  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU91/00429  
FILING DATE: 17-SEP-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: AU PK2361  
FILING DATE: 18-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, STEPHEN A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 16786/175/CHAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-987-286-24

Query Match 1.4%; Score 7; DB 1; Length 25;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAVGAA 295  
Db 12 GAAVGAA 18

RESULT 45  
US-08-425-069-38  
Sequence 38, Application US/08425069  
Patent No. 5728810  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
APPLICANT: Hinman, Michael B.  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5728810th Washington Street  
CITY: Falls Church  
STATE: Virginia

COUNTRY: U.S.A.  
ZIP: 22046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,069  
FILING DATE: 19-APR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-106P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX:  
INFORMATION FOR SEQ. ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: Internal  
ORGANISM: nephila clavipes  
FEATURE:  
NAME/KEY: peptide  
LOCATION: 1..25  
OTHER INFORMATION: /label=silk1-repeat  
US-08-425-069-38

Query Match 1.4%; Score 7; DB 1; Length 25;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGGA 300  
Db 7 AAAAGGA 13

Search completed: January 27, 2003, 16:37:39  
Job time : 41 secs



FT /note= "Mature ORP-3"  
 XX MO200144448-A2.  
 XX 21-JUN-2001.  
 XX 07-DEC-2000; 2000WO-US33158.  
 XX 16-DEC-1999; 99US-0172367.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Yue H, Lal P, Tang YT, Hillman JL, Baughn MR, Azimzal Y, Lu DAM;  
 PI WPI: 2001-390245/41.  
 DR N-PSDB; AAH24225.  
 DR Novel human oxidoreductase protein (ORP) useful for diagnosing,  
 PT treating and preventing cell proliferative, neurological, viral,  
 PT reproductive and autoimmune/inflammatory disorders associated with  
 PT abnormal expression of ORP -  
 XX Claim 1: Page 99-100; 136pp; English.  
 PS  
 XX Sequences AAB73668-AAB73694 represent 27 novel human oxidoreductase  
 CC proteins, designated ORP-1 to ORP-27 respectively, and sequences  
 CC AAH24223-AAH24249 represent cDNAs encoding ORP-1 to ORP-27. Human ORP  
 CC proteins and nucleic acids are useful for diagnosing, treating or  
 CC preventing cell proliferative disorders (e.g. arteriosclerosis,  
 CC cirrhosis, psoriasis, cancers); endocrine disorders (e.g., type I or II  
 CC diabetes mellitus, diabetes insipidus, dwarfism, hirsutism, amenorrhoea,  
 CC osteoporosis); metabolic disorders (e.g., obesity, phenylketonuria,  
 CC hypercholesterolaemia); reproductive disorders (e.g., infertility,  
 CC ovulatory and menstrual cycle defects, endometriosis, polycystic  
 CC ovary disease, disruption of spermatogenesis, impotence); neurological  
 CC disorders (e.g., epilepsy, stroke, Alzheimer's disease, Huntington's  
 CC disease, Parkinson's disease, meningitis, Creutzfeldt-Jakob disease,  
 CC cerebral palsy, muscular dystrophy, mood, anxiety and schizophrenic  
 CC disorders); viral, bacterial, fungal and parasitic infections; and  
 CC autoimmune/inflammatory disorders such as acquired immunodeficiency  
 CC syndrome (AIDS), allergies, asthma, Crohn's disease, atopic dermatitis,  
 CC gout, multiple sclerosis, rheumatoid arthritis or ulcerative colitis.  
 CC Human ORP proteins and nucleotides can be used to identify compounds  
 CC which modulate their activity or expression. ORP nucleic acid sequences  
 CC may also be used for assessing the toxicity of a test compound, to detect  
 CC upstream sequences such as promoters and regulatory elements, and to  
 CC create knock out or knock in animals or transgenic animals to model  
 CC human disease. Oligonucleotide primers derived from ORP gene sequences  
 CC may be used to detect single nucleotide polymorphisms (SNPs) and for  
 CC mapping the naturally occurring genomic sequences. Antibodies specific  
 CC for ORP proteins may be used in the diagnosis of disorders associated  
 CC with aberrant ORP expression, in assays to monitor patients being treated  
 CC with ORP or modulators thereof, and for assessing toxicity of potential  
 CC drugs.  
 CC XX  
 SQ Sequence 555 AA;  
 XX  
 Query Match 1.8%; Score 9; DB 22; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 280 GAGIAGLAA 288  
 |||||||  
 DB 31 GAGIAGLAA 39  
 RESULT 45  
 ABB58615  
 ID ABB58615 standard; Protein; 696 AA.  
 XX  
 AC ABB58615;  
 XX  
 DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2637.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW  
 XX Drosophila melanogaster.  
 OS  
 XX MO200171042-A2.  
 PN 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 PF 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PR  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-656660/75.  
 DR N-PSDB; ABL02718.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 2637; 21pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences and pharmaceutical drugs. The invention  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC XX  
 SQ Sequence 696 AA;  
 XX  
 Query Match 1.8%; Score 9; DB 22; Length 696;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 293 GAAAGGAA 301  
 |||||||  
 DB 88 GAAAGGAA 96

Search completed: January 27, 2003, 16:34:50  
 Job time : 60 secs

PR 20-OCT-2000: 2000US-0693325.  
 PR 30-NOV-2000: 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R.  
 DR WPI: 2001-476283/51.  
 DR N-PSDB: AAK52679.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20: Page 280; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activating  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SO Sequence 518 AA:  
 Query Match 1.8%; Score 9; DB 22; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 280 GAGLAGLAA 288  
 Db 37 GAGLAGLAA 45  
 IIIIIIIIII  
 RESULT 43  
 AAB12164  
 ID AAB12164 standard; Protein: 555 AA.  
 XX  
 AC AAB12164;  
 XX  
 DT 02-FEB-2001 (first entry)  
 DE  
 XX  
 DE Hydrophobic domain protein from clone HP10673 isolated from thymus cells.  
 XX  
 KW Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;  
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokineic;  
 KW hemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 PF  
 XX 17-NOV-1999: 99WO-JP06412.  
 PR 17-NOV-1998: 98JP-0326255.  
 PR 22-DEC-1998: 98JP-0364315.  
 PR 16-MAR-1999: 99JP-0069811.  
 PR 27-APR-1999: 99JP-0119299.  
 PR 19-MAY-1999: 99JP-0138169.

XX  
 PA (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 PI  
 DR WPI: 2000-387753/33.  
 DR N-PSDB: AAA62067, AAA62077.  
 XX  
 PT Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, antiinflammatory, chemokineic,  
 PT hemostatic, thrombolytic -  
 XX  
 PS Claim 1: Page 376-378; 410pp; English.  
 XX  
 CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophylaxis of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is a human protein which  
 CC has at least one hydrophobic domain. This protein may be a secretory or a  
 CC membrane protein. The present protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haematopoiesis activity, tissue growth activity,  
 CC activity/inhibin activity, chemotactic/chemokineic activity, haemostatic  
 CC inhibition activity, anti-inflammatory activity and tumour  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer.  
 XX  
 SO Sequence 555 AA:  
 Query Match 1.8%; Score 9; DB 21; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 280 GAGLAGLAA 288  
 Db 31 GAGLAGLAA 39  
 IIIIIIIIII  
 RESULT 44  
 AAB73670  
 ID AAB73670 standard; Protein: 555 AA.  
 XX  
 AC AAB73670;  
 XX  
 DT 11-SEP-2001 (first entry)  
 DE  
 XX  
 DE Human oxidoreductase protein ORP-3.  
 XX  
 KW Human oxidoreductase protein; ORP; cell proliferative disorder;  
 KW arteriosclerosis; cirrhosis; psoriasis; cancer; endocrine disorder;  
 KW diabetes mellitus; diabetes insipidus; dwarfism; hirsutism; amenorrhoea;  
 KW osteoporosis; metabolic disorder; obesity; phenylketonuria;  
 KW hypercholesterolaemia; reproductive disorder; infertility;  
 KW ovulatory defect; menstrual cycle defect; endometriosis;  
 KW polycystic ovary disease; spermatogenesis disruption; impotence;  
 KW neurological disorder; epilepsy; stroke; Alzheimer's disease;  
 KW Huntington's disease; Parkinson's disease; Creutzfeldt-Jakob disease;  
 KW schizophrenia disorder; infection; autoimmune disorder;  
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS; asthma;  
 KW allergy; Crohn's disease; atopic dermatitis; gout; multiple sclerosis;  
 KW rheumatoid arthritis; ulcerative colitis; drug screening;  
 KW toxicity screening; transgenic animal; SNP detection; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..40  
 FT /Label= signal\_peptide  
 FT Protein 41..555

OS Chlamydia trachomatis.  
 XX  
 PN MO9928475-A2.  
 XX  
 PD 10-JUN-1999.  
 XX  
 PF 27-NOV-1998: 98WO-IB01939.  
 XX  
 PR 04-NOV-1998: 98US-0107077.  
 PR 28-NOV-1997: 97FR-0015041.  
 PR 17-DEC-1997: 97FR-0016034.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Griffais R;  
 XX  
 DR WPI; 1999-371125/31.  
 XX  
 PT Genome sequence of Chlamydia trachomatis  
 XX  
 PS Disclosure; Page 1271-1272; 1755pp; English.  
 XX  
 CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
 CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
 CC can also be used to control growth of the microorganism. Chlamydia  
 CC trachomatis is responsible for a large number of diseases, e.g. eye  
 CC diseases such as conventional trachoma, nonendemic trachoma,  
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
 CC peritonitis, Bartholinitis; pneumonia in breast feeding infants;  
 CC and venereal lymphogranulomatosis. The polypeptides of the invention  
 CC may be of use in treating these diseases.  
 XX  
 SQ Sequence 481 AA;  
 Query Match 1.8%; Score 9; DB 20; Length 481;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 447 FMOQASKIA 455  
 II|IIIIII|  
 DB 446 FMOQASKIA 454  
 RESULT 41  
 AAM78562  
 ID AAM78562 standard; Protein; 513 AA.  
 XX  
 AC AAM78562;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human protein SEQ ID NO 1224.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;  
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R.  
 XX  
 DR WPI; 2001-476283/51.  
 DR N-PSDB; AAK51695.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 3481-3482; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibit activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SQ Sequence 513 AA;  
 Query Match 1.8%; Score 9; DB 22; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 280 GAGIAGLAA 288  
 II|IIIIII|  
 DB 31 GAGIAGLAA 39  
 RESULT 42  
 AAM79546  
 ID AAM79546 standard; Protein; 518 AA.  
 XX  
 AC AAM79546;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human protein SEQ ID NO 3192.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.

PR	11-AUG-1999;	99US-0148319
PR	12-AUG-1999;	99US-0148341
PR	13-AUG-1999;	99US-0148565
PR	13-AUG-1999;	99US-0148684
PR	16-AUG-1999;	99US-0149368
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149420
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149723
PR	20-AUG-1999;	99US-0149929
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150566
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	27-AUG-1999;	99US-0151080
PR	30-AUG-1999;	99US-0151303
PR	31-AUG-1999;	99US-0151348
PR	01-SEP-1999;	99US-0151930
PR	07-SEP-1999;	99US-0151930
PR	10-SEP-1999;	99US-0152363
PR	13-SEP-1999;	99US-0153070
PR	15-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	16-SEP-1999;	99US-0154039
PR	20-SEP-1999;	99US-0154779
PR	22-SEP-1999;	99US-0155139
PR	23-SEP-1999;	99US-0155486
PR	24-SEP-1999;	99US-0155659
PR	28-SEP-1999;	99US-0156458
PR	29-SEP-1999;	99US-0156586
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	07-OCT-1999;	99US-0158029
PR	08-OCT-1999;	99US-0158232
PR	12-OCT-1999;	99US-0158369
PR	13-OCT-1999;	99US-0159293
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159637
PR	18-OCT-1999;	99US-0159638
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160770
PR	21-OCT-1999;	99US-0160814
PR	22-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	25-OCT-1999;	99US-0161404
PR	25-OCT-1999;	99US-0161405
PR	26-OCT-1999;	99US-0161406
PR	26-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	28-OCT-1999;	99US-0161361
PR	28-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161992
PR	29-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162142

Query Match	1.8%	Score 9:	DB 21:	Length 382:
Best Local Similarity	100.0%	Pred. No. 18:		
Matches 9:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	280	GAGLGGIAA	288	
Db	56	GAGLGGIAA	64	

RESULT 39	
ABB92256	
ID ABB92256 standard; Protein; 477 AA.	
XX	
AC ABB92256;	
XX	
DT 31-MAY-2002 (first entry)	
XX	
DE Herbicidally active polypeptide SEQ ID NO 1467.	
XX	
KW Herbicidal; plant; agriculture; herbicide.	
XX	
OS Arabidopsis thaliana.	
XX	
PN W0200210210-A2.	
XX	
PD 07-FEB-2002.	
XX	
PF 28-AUG-2001; 2001WO-EP09892.	
XX	
PR 28-AUG-2001; 2001WO-EP09892.	
XX	
PA (FARB ) BAYER AG.	
XX	
PI Tietjen K, Weidler M;	
XX	
DR WPI; 2002-269010/31.	
XX	
PT Identifying plant target proteins for herbicidally active compounds,	
PT compiling aligning and comparing nucleic acid or amino acid sequences	
PT from plant with nucleic acid or amino acid sequences from non-plant	
PT organisms -	
XX	
PS Claim 5; SEQ ID NO 1467; 261pp + Sequence Listing; English.	
XX	
CC The invention relates to identifying target proteins	
CC (ABB90790-ABB94010) for herbicidally active compounds, comprising	
CC aligning and comparing nucleic acid or amino acid sequences from plant	
CC with nucleic acid or amino acid sequences from non-plant organisms using	
CC suitable search parameters, where plant sequences having an E-value	
CC greater by a factor of 3 than the E-value of most similar non-plant	
CC sequences are selected. The polypeptides or nucleic acids encoding them	
CC are useful for identifying modulators. The identified modulators are	
XX useful as herbicides.	
XX	
SQ Sequence 477 AA;	
Query Match 1.8%; Score 9; DB 23; Length 477;	
Best Local Similarity 100.0%; Pred. No. 22;	
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 280 GAGLAGLAA 288	
DB 53 GAGLAGLAA 61	
RESULT 40	
AAY37632	
ID AAY37632 standard; Protein; 481 AA.	
XX	
AC AAY37632;	
XX	
DT 07-OCT-1999 (first entry)	
XX	
DE Protein which is specific to Chlamydia trachomatis.	
XX	
KW Vaccine; eye disease; conventional trachoma; nongendemic trachoma;	
KW paratrachoma; inclusion conjunctivitis; genital disease; periorbitis;	
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;	
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.	
XX	

Query Match 1.8%; Score 9; DB 21; Length 379;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288  
|||||||  
DB 53 GAGLAGLAA 61

RESULT 38  
AAG10363  
ID AAG10363 standard; Protein; 382 AA.  
XX  
AC AAG10363;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8653.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP103405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137702.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144329.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144652.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-015087.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (C) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.

XX SQ Sequence 172 AA;

Query Match 1.8%; Score 9; DB 21; Length 172;  
Best Local Similarity 100.0%; Pred. No. 9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGGAGAA 304  
|||||

DB 42 AAGGAGAA 50

RESULT 36

ABG05860

ID ABG05860 standard; Protein; 354 AA.

AC ABG05860;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #5851.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW Food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-MAR-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB: AAS70047.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

PS Claim 20: SEQ ID No 36219; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 354 AA;

Query Match 1.8%; Score 9; DB 22; Length 354;  
Best Local Similarity 100.0%; Pred. No. 17;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 LAAGAAVGA 294  
|||||

DB 149 LAAGAAVGA 157

RESULT 37

AAG10364

ID AAG10364 standard; Protein; 379 AA.

AC AAG10364;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 8654.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

PF 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135253.

PR 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

CC kidney failure) and to delay or prevent the development of PKD-associated  
CC hypertension, bleeding into the cysts, pain, and renal insufficiency  
CC associated with the progression of cystic disease. In an exemplification  
CC of the invention, the expression profile of representative genes was  
CC examined in a rat model of polycystic kidney disease and compared with  
CC gene expression in normal rats. The PRBR gene (also referred to as clone  
CC PKD42) was found to be overexpressed by a factor of 2 in the rat model of  
CC PKD, while the DBI gene was underexpressed by a factor of 1.9. The  
CC present sequence represents the amino acid sequence of human PRBR.  
CC Note: The present sequence is not given in the specification although it  
CC is referred to as SEQ ID NO:3 and is stated as being given in figure 2.  
CC This sequence was obtained using the DNA sequence given in figure 2.  
XX  
XX  
SQ Sequence 169 AA;

Query Match 1.8%; Score 9; DB 21; Length 169;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310  
|||||||  
Db 117 GAAATTTVA 125

## RESULT 34

AAY93604  
ID AAY93604 standard; Protein: 169 AA.

XX  
AC AAY93604;

DT 25-SEP-2000 (first entry)

DE Protein encoded by a human peripheral benzodiazepine receptor gene.

XX  
XX Differentially expressed human gene: cardiac disease; kidney disease;  
KW inflammatory disease; I-80; prostacyclin-stimulating factor; isf-2;  
KW tissue specific mRNA; insulin-like growth factor binding protein 6;  
KW OSF-1; gas-1; YMP; BTG2; pre-B cell stimulating factor homologue; SDF1a;  
KW peripheral benzodiazepine receptor; annexin II cellular ligand; p11;  
KW congenital heart failure; dilated congestive cardiomyopathy;  
KW hypertrophic cardiomyopathy; restrictive cardiomyopathy;  
KW mitral valve disease; aortic valve disease; tricuspid valve disease;  
KW angina pectoris; myocardial infarction; cardiac arrhythmia;  
KW pulmonary hypertension; arterial hypertension; renovascular hypertension;  
KW arteriosclerosis; atherosclerosis; cardiac tumour.

XX Homo sapiens.

OS  
XX WO200035473-A2.

PN 22-JUN-2000.

PD 15-DEC-1999; 99WO-US29941.

XX 18-DEC-1998; 98US-0113008.

XX (SCIO-) SCIOS INC.

PA Stanton LW, White RT, Damm DL, Lewicki JA, Joly A, Schreiner GF;

PI WPI; 2000-451904/39.

DR N-PSDB; AAA46678.

XX Preventing, diagnosing and treating cardiac, kidney and inflammatory  
PT disorders using cardiac genes that are differentially expressed in  
PT disease states such as cardiac arrhythmia and arteriosclerosis -  
XX  
XX Disclosure: Fig 8K; 170pp; English.

XX AAY93594-y93605 are encoded by differentially expressed human genes,  
CC associated with disease states and disorders. The specification  
CC describes methods preventing, diagnosing and treating cardiac, kidney  
CC and inflammatory diseases associated with inappropriate expression of

CC differentially expressed cardiac, kidney and inflammatory genes  
CC (e.g. AAA46668-79). These genes include I-80, prostacyclin-stimulating  
CC factor, isf-2, tissue specific mRNA, insulin-like growth factor  
CC binding protein 6, OSF-1, gas-1, YMP, BTG2, pre-B cell stimulating  
CC factor homologue (SDF1a), peripheral benzodiazepine receptor, and  
CC cellular ligand of annexin II (p11), respectively. These diseases  
CC include congenital heart failure, dilated congestive cardiomyopathy,  
CC hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve  
CC disease, aortic valve disease, tricuspid valve disease, angina pectoris,  
CC myocardial infarction, cardiac arrhythmia, pulmonary hypertension,  
CC arterial hypertension, renovascular hypertension, arteriosclerosis,  
CC atherosclerosis and/or cardiac tumours.  
XX  
XX  
SQ Sequence 169 AA;

Query Match 1.8%; Score 9; DB 21; Length 169;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310  
|||||||  
Db 117 GAAATTTVA 125

## RESULT 35

AAB38041  
ID AAB38041 standard; Peptide: 172 AA.

XX  
AC AAB38041;

DT 31-JAN-2001 (first entry)

DE Fragment of human secreted protein encoded by gene 10 clone HWHGP71.

XX  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;  
KW vulnerary; anticoagulant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS  
XX WO200055371-A1.

PN 21-SEP-2000.

PD 16-MAR-2000; 2000WO-US06783.

XX 18-MAR-1999; 99US-0125055.

XX (HUMA-) HUMANA GENOME SCI INC.

PA Ruben SM, Ni J, Ebner R, Rosen CA, Shi Y, Blise C, Florence K;  
PI Komatsoulis G, Lafleur DM, Moore PA, Olsen HS, Young PE;

DR WPI; 2000-594448/56.

XX New nucleic acid molecules encoding 27 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives -  
XX  
XX Disclosure: Page 27; 453pp; English.

XX Sequences AAB37984-B38019 represent the amino acid sequences of 27  
CC human secreted proteins encoded by the genes AAC69084-C69119. The genes  
CC and proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
CC a range of human tissues disclosed in the specification. The nucleic  
CC acids, proteins, antibodies and (ant)agonists are useful in the  
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune



PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 9; DB 21; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 8.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301  
 Db 129 GAAAGGAA 137

RESULT 32  
 AAY42438  
 ID AAY42438 standard; Protein; 169 AA.  
 AC AAY42438;  
 XX  
 DT 08-DEC-1999 (first entry)  
 XX  
 DE Amino acid sequence of partial PBR from MDA-231 and MCF-7.  
 XX  
 DE  
 XX  
 KM Cancer; cell proliferation; invasive tumor; aggressive tumor; cytostatic;  
 KM antiproliferative; hypoproliferative; detection; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 OS

Key Location/Qualifiers  
 FH 1.26  
 FT Region /note="nucleotide sequence unknown"  
 FT  
 FT  
 PN WO9949316-A2.  
 XX  
 PD 30-SEP-1999.  
 XX  
 PF 25-MAR-1999; 99MO-US06515.  
 XX  
 PR 25-MAR-1999; 98US-0047652.  
 XX  
 PA (GEO) UNIV GEORGETOWN MEDICAL CENT.  
 PI Papadopoulos V, Culty M;  
 PI  
 DR MPI: 1999-580494/49.  
 DR N-PSDB: AA222636, AA222637.  
 XX  
 XX

Agents that reduce peripheral-type benzodiazepine receptor function,  
 useful for detecting and treating aggressive phenotype breast cancer  
 Clalm 32; Page 74-75; 76pp; English.

This is the amino acid sequence of the peripheral-type benzodiazepine  
 receptor, derived from the human breast cancer cell lines, MCF-7 and  
 MDA-231. The amino acid sequence differs from the wildtype PBR sequence  
 with the replacement of histidine 162 with arginine and replacement  
 of alanine 147 with a threonine.  
 CC This invention provides the means of peripheral-type benzodiazepine  
 CC receptor (PBR) modulation.  
 CC Molecules that inhibit PBR function are used to treat diseases  
 CC mediated by increased cell proliferation, particularly cancers (e.g.  
 CC cancers of the lung, stomach, prostate, ovary, testes, pancreas, liver,  
 CC cervix, vagina, skin, esophagus, nasopharynx, oropharynx, brain,  
 CC central and peripheral nervous systems, bone and cartilage), especially

CC breast cancer. PBR, or a vector encoding PBR may be used to treat  
 CC diseases mediated by reduced cell proliferation, such as developmental  
 CC retardation.  
 CC

SQ Sequence 169 AA;

Query Match 1.8%; Score 9; DB 20; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310  
 Db 117 GAAATTTVA 125

RESULT 33  
 AAB11966  
 ID AAB11966 standard; Protein; 169 AA.  
 AC AAB11966;  
 XX  
 DT 22-NOV-2000 (first entry)  
 XX  
 DE Human peripheral-type benzodiazepine receptor (PTBR).  
 XX  
 DE  
 XX  
 KM Periphrastc kidney disease; PKD; hypertension; renal failure;  
 KM polycystic kidney disease; PKD; hypertension; renal failure;  
 KM Insufficiency; autosomal dominant form; autosomal recessive form;  
 KM cystic disease; PTBR ligand; differential expression; antiproliferative;  
 KM hypotensive.  
 XX  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 PN WO200037085-A1.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 15-DEC-1999; 99MO-US29763.  
 XX  
 PR 18-DEC-1998; 98US-0113008.  
 PR 26-MAY-1999; 99US-0136208.  
 XX  
 XX  
 PA (SCIO-) SCIOS INC.  
 PA (STAN/) STANTON L W.  
 XX  
 PI Schreiner GF, Joly A, White RT;  
 PI  
 DR MPI: 2000-442524/38.  
 DR N-PSDB: AAA72048.  
 XX  
 XX  
 XX

Treatment and prevention of cystic disease, specifically polycystic  
 kidney disease, by administering ligand of peripheral benzodiazepine  
 receptor  
 Example 1; Page -: 37pp; English.

The invention relates to a method for the treatment and prevention of  
 CC diseases that involve cyst formation, particularly polycystic kidney  
 CC disease (PKD), comprising administering a ligand of a peripheral-type  
 CC benzodiazepine receptor (PTBR). The PTBR ligand can also be used for the  
 CC treatment of hypertension associated with PKD. The invention also  
 CC encompasses the diagnosis and/or prognosis of cystic disease, or the  
 CC diagnosis of a predisposition to cystic disease, via the monitoring of  
 CC changes in expression levels of PTBR or an endogenous ligand of PTBR,  
 CC such as the diazepam binding inhibitor (DBI). In PKD, the PTBR gene is  
 CC overexpressed, while the natural ligand is underexpressed. The PTBR  
 CC ligands of the invention inhibit the proliferation of epithelial cells  
 CC from PKD kidneys, i.e., they arrest cell growth but do not kill cells.  
 CC The method is used to treat diseases associated with cyst formation,  
 CC particularly both the autosomal dominant and recessive forms of PKD.  
 CC The methods can also be used for the diagnosis and prognosis of a cystic  
 CC disease or susceptibility to it. The methods may also be used to  
 CC delay or prevent cyst expansion (and therefore subsequent progression to

PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134320.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138099.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 17-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142970.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145293.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151330.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

DB 56 GAAAGCAA 64

RESULT 30

ABG41659

ID ABG41659 standard; Peptide: 153 AA.

AC ABG41659;

XX 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 31324.

XX Human: single exon probe; asthma; lung cancer; COPD; IHD;

KW Chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemostasis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

OS MO200186003-A2.

PN 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

PF 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI: 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 31324; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemostasis, pulmonary histiocytosis, lymphangioleiomyomatosis,

CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPD at

XX ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 153 AA;

OY 293 GAAAGCAA 301

DB 56 GAAAGCAA 64

Query Match 1.8%; Score 9; DB 23; Length 153;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 31

AAAG1693

ID AAAG1693 standard; Protein: 159 AA.

XX AAAG1693;

AC 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 10507.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 10507.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 23-MAR-1999; 990S-0125788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.

PR 23-APR-1999; 990S-0130891.

PR 28-APR-1999; 990S-0131491.

PR 30-APR-1999; 990S-0132048.

PR 30-APR-1999; 990S-0132407.

PR 04-MAY-1999; 990S-0132484.

PR 05-MAY-1999; 990S-0132485.

PR 06-MAY-1999; 990S-0132486.

PR 06-MAY-1999; 990S-0132487.

PR 07-MAY-1999; 990S-0132863.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-48890/53.  
DR Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
XX Example 4; SEQ ID NO: 32152; 658bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 153 AA;  
  
Query Match 1.8%; Score 9; DB 22; Length 153;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 293 GAAAGCAA 301  
DB 56 GAAAGCAA 64  
  
RESULT 28  
AAM19335  
ID AAM19335 standard; Protein; 153 AA.  
XX  
XX AAM19335;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Peptide #5769 encoded by probe for measuring cervical gene expression.  
XX  
KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
DR Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
XX Claim 27; SEQ ID NO 24161; 487bp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 153 AA;  
  
Query Match 1.8%; Score 9; DB 22; Length 153;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 293 GAAAGCAA 301  
DB 56 GAAAGCAA 64  
  
RESULT 29  
AAM32131  
ID AAM32131 standard; Protein; 153 AA.  
XX  
XX AAM32131;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #6168 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-48897/53.  
DR Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX Claim 27; SEQ ID NO 32400; 654bp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENPs:  
XX see AAI13135-AA157546). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for producing a microarray for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders.  
XX  
SQ Sequence 153 AA;  
  
Query Match 1.8%; Score 9; DB 22; Length 153;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 293 GAAAGCAA 301  
DB 56 GAAAGCAA 64

```
XX Human: gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488699/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 15; SEQ ID No 25516; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21533-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 153 AA;
XX
XX Query Match 1.8%; Score 9; DB 22; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 293 GAAAGGAA 301
XX |||||||
XX 56 GAAAGGAA 64
XX
XX Db
XX
XX RESULT 26
XX ID AAM59299 standard; Protein; 153 AA.
XX
XX AC AAM59299;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 31404.
XX
XX Human: brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX PN
```

```
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 31404; 650pp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX SQ Sequence 153 AA;
XX
XX Query Match 1.8%; Score 9; DB 22; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 293 GAAAGGAA 301
XX |||||||
XX 56 GAAAGGAA 64
XX
XX Db
XX
XX RESULT 27
XX ID AAM71846 standard; Protein; 153 AA.
XX
XX AC AAM71846;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 32152.
XX
XX Human: bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PN
```

PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149422.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 27-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0151930.  
PR 10-SEP-1999; 99US-0152363.  
PR 13-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 16-SEP-1999; 99US-0154031.  
PR 20-SEP-1999; 99US-0154039.  
PR 22-SEP-1999; 99US-0154179.  
PR 23-SEP-1999; 99US-0155139.  
PR 24-SEP-1999; 99US-0155486.  
PR 28-SEP-1999; 99US-0156559.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 9; DB 21; Length 149;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301

DB 119 GAAAGGAA 127  
|||||  
RESULT 24  
ABBS38670  
ID ABB38670 standard; Peptide; 153 AA.  
XX  
ABBS38670;  
XX  
AC 04-FEB-2002 (first entry)  
XX  
DE Peptide #6176 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
P1 Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX  
PT analyzing gene expression in human fetal liver -  
XX  
PS Claim 27; SEQ ID NO 31305; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
XX  
CC measuring human gene expression in a sample derived from human foetal  
XX  
CC liver. The single exon nucleic acid probes may be used for predicting,  
XX  
CC measuring and displaying gene expression in samples derived from human  
XX  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
XX  
CC nucleic acid probe of the invention.  
XX  
CC Note: The sequence data for this patent did not form part of the  
XX  
CC printed specification, but was obtained in electronic format directly  
XX  
XX from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
SQ Sequence 153 AA:  
QY 293 GAAAGGAA 301  
|||||  
Db 56 GAAAGGAA 64  
RESULT 25  
ABBS3746  
ID ABB23746 standard; Protein; 153 AA.  
XX  
AC ABB23746;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #5745 encoded by probe for measuring heart cell gene expression.

PR 29-OCT-1999; 99US-0162142.  
Query Match 1.8%; Score 9; DB 21; Length 118;  
Best Local Similarity 100.0%; Fred. No. 6.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 293 GAAAGCAA 301  
| | | | | | | |  
DB 88 GAAAGCAA 96  
RESULT 23  
AAG11694  
ID AAG11694 standard; Protein; 149 AA.  
XX  
AC AAG11694;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10508.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
OS  
PN EPI03405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 200EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130409.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145152.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.

PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154779.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.



PR 05-FEB-1999; 99GB-0002555.  
XX (NEUT-) NEUTECH PHARMA PLC.  
XX  
XX Burnie JP, Matthews RC;  
XX  
XX WPI: 2000-543485/49.  
XX  
XX  
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
PT preventing and treating C. pneumoniae infection and atherosclerosis,  
PT including coronary atherosclerosis -  
XX  
XX Claim 5; Page 34; 35pp; English.  
XX  
XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae  
CC protein. The protein, immunogenic fragments of it, nucleotide sequences  
CC encoding it, or inhibitor specific against it are used to manufacturing  
CC a medicament for the treatment of infection due to C. pneumoniae.  
CC An antibody specific against the protein can diagnose a C. pneumoniae  
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,  
CC including coronary atherosclerosis, caused by C. pneumoniae can also  
CC be prevented or treated.  
XX  
XX  
SQ Sequence 9 AA;  
Query Match 1.8%; Score 9; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 235 EYDEMKAEE 243  
Db 1 EYDEMKAEE 9  
RESULT 21  
AAR72669  
ID AAR72669 standard; Protein: 111 AA.  
XX  
XX AAR72669;  
AC  
DT 25-OCT-1995 (first entry)  
XX  
XX Cladosporium herbarum allergen Clah1.  
XX  
XX Cladosporium herbarum allergen Clah1.  
XX  
XX Fungal spore; allergen; Clah1; allergy; ribosomal protein. RL42.  
XX  
XX Cladosporium herbarum.  
XX  
XX W09506121-A.  
XX  
XX 02-MAR-1995.  
XX  
XX 24-AUG-1994; 94MO-AT00120.  
XX  
XX 27-AUG-1993; 93AT-0001725.  
XX  
XX (BIOM-) BIOMAY PRODN & HANDELSGES MBH.  
XX  
XX Achatz G, Breitenbach M, Ebner C, Hirschwehr R;  
PI Kraft D, Lechenauer E, Oberkofler H, Prillinger H;  
PI Simon B, Unger A;  
XX  
XX WPI: 1995-106850/14.  
XX  
XX N-PSDB; AA087846.  
XX  
XX Allergens derived from Cladosporium herbarum spores - also  
PT recombinant DNA for expressing the allergens, useful for in vitro  
PT allergy detection  
XX  
XX Claim 1; Page 25; 35pp; German.  
XX  
XX Spores of Cladosporium herbarum are the most common fungal spores  
CC found in the air; they can cause allergic reactions. Various Clah  
CC

CC allergens and sequences encoding them have now been isolated. The  
CC mature Clah1 allergen has mol. wt. 11 kD and is encoded by cDNA  
CC sequence AA087846. The allergen has homology to the ribosomal protein  
CC RL42. Potential epitopic subfragments were identified by computer  
CC analysis of the amino acid sequence. See AAR72670-R72674 for potential  
CC B-cell epitopes and AAR72675-R72677 for potential T-cell epitopes.  
XX  
XX  
SQ Sequence 111 AA;  
Query Match 1.8%; Score 9; DB 16; Length 111;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 293 GAAAGCAA 301  
Db 77 GAAAGCAA 85  
RESULT 22  
AAG11695  
ID AAG11695 standard; Protein: 118 AA.  
XX  
XX AAG11695;  
AC  
DT 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10509.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 9905-0121825.  
XX  
XX 05-MAR-1999; 9905-0123180.  
XX  
XX 09-MAR-1999; 9905-0123548.  
XX  
XX 23-MAR-1999; 9905-0125788.  
XX  
XX 25-MAR-1999; 9905-0126264.  
XX  
XX 29-MAR-1999; 9905-0126785.  
XX  
XX 01-APR-1999; 9905-0127462.  
XX  
XX 06-APR-1999; 9905-0128234.  
XX  
XX 08-APR-1999; 9905-0128714.  
XX  
XX 16-APR-1999; 9905-0129845.  
XX  
XX 19-APR-1999; 9905-0130077.  
XX  
XX 21-APR-1999; 9905-0130449.  
XX  
XX 23-APR-1999; 9905-0130510.  
XX  
XX 23-APR-1999; 9905-0130891.  
XX  
XX 28-APR-1999; 9905-0131449.  
XX  
XX 30-APR-1999; 9905-0132048.  
XX  
XX 30-APR-1999; 9905-0132407.  
XX  
XX 04-MAY-1999; 9905-0132484.  
XX  
XX 05-MAY-1999; 9905-0132485.  
XX  
XX 06-MAY-1999; 9905-0132486.  
XX  
XX 06-MAY-1999; 9905-0132487.  
XX  
XX 07-MAY-1999; 9905-0132863.  
XX  
XX 11-MAY-1999; 9905-0134256.  
XX  
XX 14-MAY-1999; 9905-0134218.  
XX  
XX 14-MAY-1999; 9905-0134219.  
XX  
XX 14-MAY-1999; 9905-0134221.  
XX  
XX 14-MAY-1999; 9905-0134370.  
XX  
XX 18-MAY-1999; 9905-0134768.  
XX  
XX 19-MAY-1999; 9905-0134941.  
XX  
XX 20-MAY-1999; 9905-0135124.  
XX  
XX 21-MAY-1999; 9905-0135353.  
XX  
XX 24-MAY-1999; 9905-0135629.  
XX  
XX 25-MAY-1999; 9905-0136021.  
XX  
XX

CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,  
 CC including coronary atherosclerosis, caused by C. pneumoniae can also  
 CC be prevented or treated.  
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 175 GVTPESEVI 183  
 DB 1 GVTPESEVI 9

RESULT 18  
 AAB08330

ID AAB08330 standard; Peptide: 9 AA.

XX AAB08330;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis;  
 KM coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

XX 28-JAN-2000; 2000WO-GB00237.

XX 05-FEB-1999; 99GB-0002555.

XX (NEUT-) NEUTEC PHARMA PLC.

XX Burnie JP, Mathews RC;

XX WPI: 2000-543485/49.

XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
 PT preventing and treating C. pneumoniae infection and atherosclerosis,  
 PT including coronary atherosclerosis -  
 XX

XX Claim 5; Page 34; 35pp; English.

XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae  
 CC protein. The protein, immunogenic fragments of it, nucleotide sequences  
 CC encoding it, or inhibitor specific against it are used to manufacturing  
 CC a medicament for the treatment of infection due to C. pneumoniae.  
 CC An antibody specific against the protein can diagnose a C. pneumoniae  
 CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,  
 CC including coronary atherosclerosis, caused by C. pneumoniae can also  
 CC be prevented or treated.  
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 37 SGNETKQIQ 45  
 DB 1 SGNETKQIQ 9

RESULT 19  
 AAB08333  
 ID AAB08333 standard; Peptide: 9 AA.

XX

AAB08333;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis;  
 KM coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

XX 28-JAN-2000; 2000WO-GB00237.

XX 05-FEB-1999; 99GB-0002555.

XX (NEUT-) NEUTEC PHARMA PLC.

XX Burnie JP, Mathews RC;

XX WPI: 2000-543485/49.

XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
 PT preventing and treating C. pneumoniae infection and atherosclerosis,  
 PT including coronary atherosclerosis -  
 XX

XX Claim 5; Page 34; 35pp; English.

XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae  
 CC protein. The protein, immunogenic fragments of it, nucleotide sequences  
 CC encoding it, or inhibitor specific against it are used to manufacturing  
 CC a medicament for the treatment of infection due to C. pneumoniae.  
 CC An antibody specific against the protein can diagnose a C. pneumoniae  
 CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,  
 CC including coronary atherosclerosis, caused by C. pneumoniae can also  
 CC be prevented or treated.  
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 SKSMESTLE 135  
 DB 1 SKSMESTLE 9

RESULT 20  
 AAB08334

ID AAB08334 standard; Peptide: 9 AA.

XX AAB08334;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis;  
 KM coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

XX 28-JAN-2000; 2000WO-GB00237.

XX

XX 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skelky YAW, Persing DH, Mitcham JL, Mang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI: 2001-616774/71.  
DR N-PSDB: AAS59546.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -  
XX  
PS Example 1; SEQ ID No 11250; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence. For example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 755 AA:  
  
Query Match 2.0%; Score 10; DB 22; Length 755;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 281 AGLAGLAAGA 290  
          |||||  
DB 575 AGLAGLAAGA 584  
  
RESULT 16  
AAB08324  
ID AAB08324 standard; Peptide; 9 AA.  
XX  
AC AAB08324;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE Epitope derived from a hlamydia pneumoniae protein.  
XX  
KM Chlamydia pneumoniae protein; infection; atherosclerosis;  
KM coronary atherosclerosis; epitope.  
XX  
OS Chlamydia pneumoniae.  
XX  
PN WO200046359-A2.  
XX  
PD 10-AUG-2000.  
XX  
PF 28-JAN-2000; 2000WO-GB00237.  
XX  
PR 05-FEB-1999; 99GB-0002555.  
XX

XX (NEUT-) NEUTEC PHARMA PLC.  
PA  
XX  
PI Burnie JP, Matthews RC;  
XX  
DR WPI: 2000-543485/49.  
XX  
XX  
PT New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
PT preventing and treating C. pneumoniae infection and atherosclerosis,  
PT including coronary atherosclerosis -  
XX  
PS Claim 5; Page 32; 35pp; English.  
XX  
CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae  
CC protein. The protein, immunogenic fragments of it, nucleotide sequences  
CC encoding it, or inhibitor specific against it are used to manufacturing  
CC a medicament for the treatment of infection due to C. pneumoniae.  
CC An antibody specific against the protein can diagnose a C. pneumoniae  
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,  
CC including coronary atherosclerosis, caused by C. pneumoniae can also  
CC be prevented or treated.  
XX  
SQ Sequence 9 AA:  
  
Query Match 1.8%; Score 9; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 163 SAKLEPEL 171  
          |||||  
DB 1 SAKLEPEL 9  
  
RESULT 17  
AAB08326  
ID AAB08326 standard; Peptide; 9 AA.  
XX  
AC AAB08326;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE Epitope derived from a hlamydia pneumoniae protein.  
XX  
KM Chlamydia pneumoniae protein; infection; atherosclerosis;  
KM coronary atherosclerosis; epitope.  
XX  
OS Chlamydia pneumoniae.  
XX  
PN WO200046359-A2.  
XX  
PD 10-AUG-2000.  
XX  
PF 28-JAN-2000; 2000WO-GB00237.  
XX  
PR 05-FEB-1999; 99GB-0002555.  
XX  
PA (NEUT-) NEUTEC PHARMA PLC.  
XX  
PI Burnie JP, Matthews RC;  
XX  
DR WPI: 2000-543485/49.  
XX  
PT New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
PT preventing and treating C. pneumoniae infection and atherosclerosis,  
PT including coronary atherosclerosis -  
XX  
PS Claim 5; Page 33; 35pp; English.  
XX  
CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae  
CC protein. The protein, immunogenic fragments of it, nucleotide sequences  
CC encoding it, or inhibitor specific against it are used to manufacturing  
CC a medicament for the treatment of infection due to C. pneumoniae.  
CC An antibody specific against the protein can diagnose a C. pneumoniae  
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,  
CC including coronary atherosclerosis, caused by C. pneumoniae can also  
CC be prevented or treated.

[illegible]

OS	Unidentified.
XX	
PN	US2002052307-A1.
PD	
XX	
PF	02-MAY-2002.
XX	
PI	08-JAN-2001; 2001US-0756071.
XX	
PR	07-JAN-2000; 2000US-175005P.
PR	04-OCT-1994; 94US-0317450.
PR	18-FEB-1997; 97US-0800593.
PR	15-SEP-2000; 2000US-0663147.
XX	
PA	(TRYG/) TRYGVASON K.
PA	(KALU/) KALIUNKI P.
XX	(PYKE/) PYKE C.
PI	Trygvason K, Kalunki P, Pyke C;
XX	
DR	WPI: 2002-434824/46.
XX	
PT	Modulating laminin 5 gamma 2 chain interactions of invasive carcinogens
PT	for treating cancers and promoting attachment of cultured cells in
PT	vitro -
XX	
PS	Disclosure; Page 34-36; 51pp; English.
XX	
CC	The invention comprises a method of inhibiting the laminin gamma-2 chain
CC	interactions of invasive carcinogens with surrounding tissues - by using
CC	anti-gamma-2 chain antibodies to inhibit the gamma-2 chain biological
CC	activity of the invasive carcinogens. The invention also comprises a
CC	method for promoting adhesion of epithelial cells by exposing the cells
CC	to intact laminin-5 molecules. The first method of the invention is
CC	useful for preventing gamma 2 chain interactions of invasive carcinogens
CC	with surrounding tissues. The second method of the invention is useful
CC	for promoting adhesion of cultured epithelial cells.
CC	NOTE: The present sequence is shown as a protein in the sequence listing,
CC	however in a figure of the invention this sequence is shown to be a DNA
CC	sequence (see AA42916).
XX	
SQ	Sequence 720 AA;
	Query Match 2.0%; Score 10; DB 23; Length 720;
	Best Local Similarity 100.0%; Pred. No. 4.1;
	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	293 GAAGAAGGAAG 302
DB	602 GAAGAAGGAAG 611
RESULT 15	
ID	AAU50055 standard; Protein: 755 AA.
XX	
AC	AAU50055;
XX	
DT	27-FEB-2002 (first entry)
XX	
DE	Propionibacterium acnes immunogenic protein #10951.
XX	
KM	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KV	uvellitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW	Inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KX	dematological; osteopathic; neuroprotectant.
XX	
OS	Propionibacterium acnes.
XX	
PN	WO200181581-A2.
XX	
PD	01-NOV-2001.
XX	
PF	20-APR-2001; 2001WO-US12865.

```

AC AAB08336;
XX
DF 04-DEC-2000 (first entry)
XX
DE Epitope derived from a hlamydia pneumoniae protein.
XX
XX Chlamydia pneumoniae protein; infection; atherosclerosis;
KW coronary atherosclerosis; epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO200046359-A2.
XX
PD 10-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00237.
XX
PR 05-FEB-1999; 99GB-0002555.
XX
PA (NEUT-) NEUTEC PHARMA PLC.
XX
PI Burnie JP, Matthews RC;
XX
DR WPI; 2000-543485/49.
XX
PR New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
PT preventing and treating C. pneumoniae infection and atherosclerosis,
PT including coronary atherosclerosis -
XX
XX Claim 5; Page 35; 35pp; English.
PS
XX AAB08334-36 represent epitopes derived from a Chlamydia pneumoniae
CC protein. The protein, immunogenic fragments of it, nucleotide sequences
CC encoding it, or inhibitor specific against it are used to manufacturing
CC a medicament for the treatment of infection due to C. pneumoniae.
CC An antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also
CC be prevented or treated.
XX
SQ Sequence 14 AA;
Query Match 2.8%; Score 14; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 167 ETPELPKPGVTPRS 180
DB 1 ETPELPKPGVTPRS 14
RESULF 12
AAB58985
ID AAB58985 standard; Protein; 1937 AA.
XX
XX ABB58985;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 3747.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
PR 23-MAR-2000; 2000US-191637P.

```

```

PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL03088.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 3747; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1937 AA;
Query Match 2.2%; Score 11; DB 22; Length 1937;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 294 AAAAGCAAGCA 304
DB 343 AAAAGCAAGCA 353
RESULF 13
ABP31472
ID ABP31472 standard; Protein; 129 AA.
XX
XX ABP31472;
AC
XX 09-JUL-2002 (first entry)
DT
XX Human ORF445 protein, SEQ ID NO:890.
DE
XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnerability;
KW vasotropic; antipsoriatic; antidiabetic; cytosolic; nocitropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US17076.
PR 24-MAY-2000; 2000US-206690P.
PA (CURA-) CURAGEN CORP.

```

```

AC AAR94585;
XX
XX 11-NOV-1996 (first entry)
XX
XX DHFR/C. pneumoniae antigen variant fusion protein.
XX
XX Polypeptide antigen: strain YK41; plasmid; probe;
XX PCPN533T; primer; assay; detection; antibody; diagnosis;
XX infection; fusion protein; dihydrofolate reductase; DHFR;
XX variant; Chlamydia pneumoniae.
XX
XX Synthetic.
XX
XX Key
XX Protein 1..160
XX FT /label= dihydrofolate reductase
XX FT 162..432
XX FT /label= C. pneumoniae antigen
XX
XX WO9609320-A1.
XX
XX 28-MAR-1996.
XX
XX 20-SEP-1995; 95WO-JP01896.
XX
XX 28-APR-1995; 95JP-0106011.
XX 20-SEP-1994; 94JP-0224711.
XX 28-APR-1995; 95JP-0106006.
XX 28-APR-1995; 95JP-0106008.
XX 28-APR-1995; 95JP-0106009.
XX 28-APR-1995; 95JP-0106010.
XX
XX (HITB ) HITACHI CHEM CO LTD.
XX
XX Izutsu H, Matsumoto A, Obara K;
XX WPI; 1996-188399/19.
XX DR N-PSDB; AAT14619.
XX
XX Recombinant Chlamydia pneumoniae antigen and antibodies to it -
XX used for detection and assay of C. pneumoniae e.g. in clinical
XX diagnosis
XX
XX Claim 23; Pages 103-107; 128pp; Japanese.
XX
XX The present sequence is the dihydrofolate reductase (DHFR)/
XX C. pneumoniae polypeptide antigen variant fusion protein.
XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted
XX to prep. a lambda gt11 DNA library. The library was then screened
XX with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
XX fusing spleen cells from a mouse infected with YK41 with myeloma
XX p3/NS1/-A44-1 to produce a Mab expressing hybridoma. The DNA
XX obtd. was then fused with DHFR DNA and the expression vector
XX PAD431 to give PCPN533T. The plasmid was used to transform an
XX E. coli host, which was cultured to give an antigenic polypeptide
XX fusion protein. The fusion protein and primers and probes derived
XX from its DNA can be used in assays for the detection of the
XX antigenic polypeptide antibodies and DNA, respectively, useful in
XX the diagnosis of C. pneumoniae infection.
XX
XX Sequence 432 AA:
XX
XX Query Match 52.2%; Score 259; DB 17; Length 432;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-222;
XX Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 4 MSISSSGPDNKNIMSOVLSTPGQVPOQDKLSGNETKQIQOTRGKNTMESDATTAG 63
XX |||||||
XX DB 162 MSISSSGPDNKNIMSOVLSTPGQVPOQDKLSGNETKQIQOTRGKNTMESDATTAG 221
XX |||||||
XX QY 64 ASGKDKTSTTETETAPQOGVAAGKSSSSQKAGADTGVSGAAATATATKIAMQTSI 123
XX |||||||
XX DB 222 ASGKDKTSTTETETAPQOGVAAGKSSSSQKAGADTGVSGAAATATATKIAMQTSI 281

```

```

QY 124 EASKSMESTLESTIOSISAQOMEVEAVVVAALSGKSSGSAKETPELPKPGVTPRSEVI 183
XX |||||||
XX DB 282 EASKSMESTLESTIOSISAQOMEVEAVVVAALSGKSSGSAKETPELPKPGVTPRSEVI 341
XX |||||||
XX QY 184 EIGIALAKAIQITGEATKRSALSNVASTOAOADOTNKLGEKQAIKIDKEEREYQEKKA 243
XX |||||||
XX DB 342 EIGIALAKAIQITGEATKRSALSNVASTOAOADOTNKLGEKQAIKIDKEEREYQEKKA 401
XX |||||||
XX QY 244 QKSKDLEGTMDFVNTVMAIA 262
XX |||||||
XX DB 402 QKSKDLEGTMDFVNTVMAIA 420
XX
XX RESULT 10
XX AAB08335
XX ID AAB08335 standard; Peptide; 14 AA.
XX
XX AC AAB08335;
XX
XX DT 04-DEC-2000 (first entry)
XX
XX DE Epitope derived from a hlamydia pneumoniae protein.
XX
XX KM Chlamydia pneumoniae protein: infection; atherosclerosis;
XX coronary atherosclerosis; epitope.
XX
XX OS Chlamydia pneumoniae.
XX
XX PN WO200046359-A2.
XX
XX PD 10-AUG-2000.
XX
XX XX 28-JAN-2000; 2000WO-GB00237.
XX
XX PF 05-FEB-1999; 99GB-0002555.
XX
XX PR (NEUT-) NEUTEC PHARMA PLC.
XX
XX PA Burnie JP, Matthews RC;
XX
XX PI WPI; 2000-543485/49.
XX
XX DR New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and atherosclerosis,
XX including coronary atherosclerosis -
XX
XX PS Claim 5; Page 34; 35pp; English.
XX
XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
XX protein. The protein, immunogenic fragments of it, nucleotide sequences
XX encoding it, or inhibitor specific against it are used to manufacturing
XX a medicament for the treatment of infection due to C. pneumoniae.
XX An antibody specific against the protein can diagnose a C. pneumoniae
XX infection. C. pneumoniae infection can be prevented. Atherosclerosis,
XX including coronary atherosclerosis, caused by C. pneumoniae can also
XX be prevented or treated.
XX
XX Sequence 14 AA:
XX
XX Query Match 2.8%; Score 14; DB 21; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 223 EKQAIKIDKEEREY 236
XX |||||||
XX DB 1 EKQAIKIDKEEREY 14
XX |||||||
XX
XX RESULT 11
XX AAB08336
XX ID AAB08336 standard; Peptide; 14 AA.
XX

```

```
XX 20-SEP-1995; 95WO-JP01896.
PF 28-APR-1995; 95JP-0106011.
XX 20-SEP-1994; 94JP-0224711.
PR 20-SEP-1994; 94JP-0224711.
XX 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
XX 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX (HITB ) HITACHI CHEM CO LTD.
PA Izutsu H, Matsumoto A, Obara K;
XX WPI: 1996-188399/19.
XX N-PSDB; AAT14613.
DR Recombinant Chlamydia pneumoniae antigen and antibodies to it
XX used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX Claim 2; Pages 64-66; 128pp; Japanese.
XX The present sequence is a variant of the C. pneumoniae
CC polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gtl1 DNA
CC library. The library was then screened with an anti-YK41
CC monoclonal antibody (Mab), which was prepd. by fusing spleen cells
CC from a mouse infected with YK41 with myeloma P3/NS1/1-A94-1 to
CC produce a Mab expressing hybridoma. The DNA obtd. was then fused
CC with the expression vector pAD431 to give pCPMS3alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.
CC Sequence 271 AA:
SQ
Query Match 52.2%; Score 259; DB 17; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.3e-222;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKIOIOTROGKNTMESDATTAG 63
DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKIOIOTROGKNTMESDATTAG 60
OY 64 ASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTAATATKIAMQTSI 123
DB 61 ASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTAATATKIAMQTSI 120
OY 124 BEASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSAALETPPELPKPGVTRSEYI 183
DB 121 BEASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSAALETPPELPKPGVTRSEYI 180
OY 184 EIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLIGKOAIRIDKREERYOEMKAAE 243
DB 181 EIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLIGKOAIRIDKREERYOEMKAAE 240
OY 244 OKSKDLEGTMDTVNTVMIA 262
DB 241 OKSKDLEGTMDTVNTVMIA 259
RESULT 8
AAB08323
ID AAB08323 standard; Protein: 302 AA.
XX
XX AAB08323;
XX
XX 04-DEC-2000 (first entry)
XX Amino acid sequence of a Chlamydia pneumoniae His-tagged protein.
DE
```

```
XX Chlamydia pneumoniae protein; infection; atherosclerosis;
KW coronary atherosclerosis.
XX Synthetic.
XX Chlamydia pneumoniae.
OS Key Location/Qualifiers
XX Cleavage-site 1..30
XX FT /note= "S-tag and thrombin cleavage site"
XX FT Misc-difference 292..302
XX FT /note= "Histidine tag"
XX WO200046359-A2.
XX 10-AUG-2000.
XX 28-JAN-2000; 2000WO-GB00237.
XX 05-FEB-1999; 99GB-0002555.
XX (NEUT-) NEUTEC PHARMA PLC.
XX Burnie JP, Matthews RC;
XX WPI: 2000-543485/49.
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and atherosclerosis,
XX including coronary atherosclerosis -
XX Example; Page 31-32; 35pp; English.
XX The present sequence represents a His-tagged Chlamydia pneumoniae
XX protein. The protein, immunogenic fragments of it, nucleotide sequences
XX encoding it, or inhibitor specific against it are used to manufacturing
XX a medicament for the treatment of infection due to C. pneumoniae.
XX An antibody specific against the protein can diagnose a C. pneumoniae
XX infection. C. pneumoniae infection can be prevented. Atherosclerosis,
XX including coronary atherosclerosis, caused by C. pneumoniae can also
XX be prevented or treated.
XX Sequence 302 AA:
SQ
Query Match 52.2%; Score 259; DB 21; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.7e-222;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DTNMSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKIOIOTROGKNTMESDATTAG 60
DB 33 DTNMSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKIOIOTROGKNTMESDATTAG 92
OY 61 IAGASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTAATATKIAMQ 120
DB 93 IAGASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTAATATKIAMQ 152
OY 121 TSIEFASKSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSAALETPPELPKPGVTRPS 180
DB 153 TSIEFASKSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSAALETPPELPKPGVTRPS 212
OY 181 EYIEIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLIGKOAIRIDKREERYOEMK 240
DB 213 EYIEIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLIGKOAIRIDKREERYOEMK 272
OY 241 AAEQSKDLEGTMDTVNTV 259
DB 273 AAEQSKDLEGTMDTVNTV 291
RESULT 9
AAR94585
ID AAR94585 standard; Protein: 432 AA.
XX
```

PF 20-SEP-1995; 95WO-JP01896.  
 XX 28-APR-1995; 95JP-0106011.  
 PR 20-SEP-1994; 94JP-0224711.  
 PR 28-APR-1995; 95JP-0106006.  
 PR 28-APR-1995; 95JP-0106008.  
 PR 28-APR-1995; 95JP-0106009.  
 PR 28-APR-1995; 95JP-0106010.  
 XX (HITB ) HITACHI CHEM CO LTD.  
 XX Izutsu H, Matsumoto A, Obara K;  
 DR WPI: 1996-188399/19.  
 DR N-PSDB; AAT14622.  
 XX Recombinant Chlamydia pneumoniae antigen and antibodies to it  
 PT used for detection and assay of C. pneumoniae e.g. in clinical  
 PT diagnosis  
 XX  
 XX Example 1; Pages 75-77; 128pp; Japanese.  
 PS  
 XX The present sequence is the C. pneumoniae polypeptide antigen  
 CC polypeptide A clone, 53-35. C. pneumoniae strain YK41 was  
 CC cultured and genomic DNA extracted to prep. a lambda g11 DNA  
 CC library. The library was then screened with an anti-YK41  
 CC monoclonal antibody (Mab), which was prepd. by fusing spleen cells  
 CC from a mouse infected with YK41 with myeloma P3/NS1/7-4g4-1 to  
 CC produce a Mab expressing hybridoma. The DNA cdt. was then fused  
 CC with the expression vector pBD431 to give pcPN533alpha. The  
 CC plasmid was used to transform an E. coli host, which was cultured  
 CC to give the antigenic polypeptide, polypeptide A. Polypeptide A  
 CC and primers and probes derived from its DNA can be used in assays  
 CC for the detection of polypeptide A antibodies and DNA,  
 CC respectively, useful in the diagnosis of C. pneumoniae infection.  
 XX  
 XX Sequence 259 AA;  
 SQ

Query Match 52.2%; Score 259; DB 17; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-222;  
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSOVLSTPGVPOQDKLSGNETKOIQOTROGKNTMESDPTIAG 63  
 DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOQDKLSGNETKOIQOTROGKNTMESDPTIAG 60  
 QY 64 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSAAATTAATKIAMQTSI 123  
 DB 61 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSAAATTAATKIAMQTSI 120  
 QY 124 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 183  
 DB 121 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 180  
 QY 184 EIGLALAKAIOTLGEATKSAISNASTOAOADQTNKLGLEKQAIKIDKEREYQEMKAAE 243  
 DB 181 EIGLALAKAIOTLGEATKSAISNASTOAOADQTNKLGLEKQAIKIDKEREYQEMKAAE 240  
 QY 244 OKSKDLEGTMDTVNTVMA 262  
 DB 241 OKSKDLEGTMDTVNTVMA 259

RESULT 6  
 AAW01743  
 ID AAW01743 standard; Protein; 259 AA.  
 AC AAW01743;  
 XX  
 XX 22-APR-1997 (first entry)  
 DE C. pneumoniae 53 kDa antigen.  
 XX

KW antigen; antibody; detection; determination; epitope.  
 XX Chlamydia pneumoniae.  
 OS JP08304404-A.  
 XX  
 PN 22-NOV-1996.  
 PD  
 XX 28-APR-1995; 95JP-0106014.  
 PF 28-APR-1995; 95JP-0106014.  
 XX 28-APR-1995; 95JP-0106014.  
 PR (HITB ) HITACHI CHEM CO LTD.  
 XX  
 XX WPI: 1997-056178/06.  
 DR N-PSDB; AAT59311.  
 XX  
 XX Detection and determination of anti-Chlamydia pneumoniae antibody -  
 PT using the polypeptide C as the antigen  
 PT  
 XX  
 XX Example 3; Page 17-18; 18pp; Japanese.  
 PS  
 XX The protein is that of the Chlamydia pneumoniae 53 kDa antigen. A  
 CC method for the detection and determination of anti-C. pneumoniae  
 CC antibodies in a sample comprises using at least 5 consecutive amino  
 CC acids of the polypeptide C 73 kDa antigen (AAW01742).  
 CC  
 XX  
 XX Sequence 259 AA;  
 SQ

Query Match 52.2%; Score 259; DB 18; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-222;  
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSOVLSTPGVPOQDKLSGNETKOIQOTROGKNTMESDPTIAG 63  
 DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOQDKLSGNETKOIQOTROGKNTMESDPTIAG 60  
 QY 64 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSAAATTAATKIAMQTSI 123  
 DB 61 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSAAATTAATKIAMQTSI 120  
 QY 124 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 183  
 DB 121 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 180  
 QY 184 EIGLALAKAIOTLGEATKSAISNASTOAOADQTNKLGLEKQAIKIDKEREYQEMKAAE 243  
 DB 181 EIGLALAKAIOTLGEATKSAISNASTOAOADQTNKLGLEKQAIKIDKEREYQEMKAAE 240  
 QY 244 OKSKDLEGTMDTVNTVMA 262  
 DB 241 OKSKDLEGTMDTVNTVMA 259

RESULT 7  
 AAR94580  
 ID AAR94580 standard; Protein; 271 AA.  
 AC AAR94580;  
 XX  
 XX 07-NOV-1996 (first entry)  
 DE C. pneumoniae polypeptide antigen (polypeptide A) variant.  
 XX  
 XX Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;  
 KW pcPN533alpha; primer; assay; detection; antibody; diagnosis;  
 XX infection; variant.  
 XX Chlamydia pneumoniae.  
 OS  
 XX W09609320-A1.  
 PN  
 XX 28-MAR-1996.  
 PD



Matches 486; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 4 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63
Db 162 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 221
OY 64 ASGDKTSTTETETAPQGVAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 123
Db 222 ASGDKTSTTETETAPQGVAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 281
OY 124 EBASKSMESTLESLSAQAQKEVEAVVAALSGKSSGSAAKLETPPEPKPGVTPRSEVI 183
Db 282 EBASKSMESTLESLSAQAQKEVEAVVAALSGKSSGSAAKLETPPEPKPGVTPRSEVI 341
OY 184 ETGLALAKAIOTLGEATKSALSNASTQAADQTNKLGLEKAIKIDKEREYOEKMAE 243
Db 342 ETGLALAKAIOTLGEATKSALSNASTQAADQTNKLGLEKAIKIDKEREYOEKMAE 401
OY 244 QSKDLEGTMDTVNTVMTAVSAIVTIVISIVAAIFTCGAGLAGLAAGAAGAAAGAGA 303
Db 402 QSKDLEGTMDTVNTVMTAVSAIVTIVISIVAAIFTCGAGLAGLAAGAAGAAAGAGA 461
OY 304 AAATTVATQITVQAVVQAVQAVTAVRQATTAIKAAVKSIGKAFITLVKAIAKAIK 363
Db 462 AAATTVATQITVQAVVQAVQAVTAVRQATTAIKAAVKSIGKAFITLVKAIAKAIK 521
OY 364 GISKVFAGTOMIAKNPKLSKVISLTSKWTGVGVVVAAPALGKGMQJLSEMOON 423
Db 522 GISKVFAGTOMIAKNPKLSKVISLTSKWTGVGVVVAAPALGKGMQJLSEMOON 581
OY 424 VAOQFQEVGKIQAAADMTSMFTQFMQAASKIASKQTESNEMTQKATKLGAIILKAYAI 483
Db 582 VAOQFQEVGKIQAAADMTSMFTQFMQAASKIASKQTESNEMTQKATKLGAIILKAYAI 641
OY 484 SGAIAGA 490
Db 642 SGAIAGA 648

```

RESULT 4  
AAR94579  
ID AAR94579 standard; Protein: 488 AA.  
AC AAR94579;  
XX  
DT 07-NOV-1996 (first entry)  
XX  
DE Chlamydia pneumoniae polypeptide antigen (polypeptide A).  
XX  
KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;  
KW PCPN533alpha; primer; assay; detection; antibody; diagnosis;  
KW Infection.  
XX  
OS Chlamydia pneumoniae.  
XX  
XX  
XX Key Location/Qualifiers  
XX Peptide 1.259  
XX /note= "claimed N-terminal fragment"  
XX  
XX WO9609320-A1.  
XX  
XX PD 28-MAR-1996.  
XX  
XX PF 20-SEP-1995; 95WO-JP01896.  
XX  
XX PR 28-APR-1995; 95JP-0106011.  
XX PR 20-SEP-1994; 94JP-0224711.  
XX PR 28-APR-1995; 95JP-0106006.  
XX PR 28-APR-1995; 95JP-0106008.  
XX PR 28-APR-1995; 95JP-0106009.  
XX PR 28-APR-1995; 95JP-0106010.  
XX  
XX PA (HITB ) HITACHI CHEM CO LTD.

XX  
PI Izutsu H, Matsumoto A, Obara K;  
XX  
DR WPI: 1996-188399/19.  
DR N-PSDB: AAT14612.  
XX  
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it  
PT used for detection and assay of C. pneumoniae e.g. in clinical  
PT diagnosis  
XX  
PS Claim 1: Pages 60-64; 128pp; Japanese.  
XX  
XX The present sequence is the C. pneumoniae polypeptide antigen,  
CC polypeptide A. C. pneumoniae strain YK41 was cultured and genomic  
CC DNA extracted to prep. a lambda gt11 DNA library. The library was  
CC then screened with an anti-YK41 monoclonal antibody (MAb), which  
CC was prep. by fusing spleen cells from a mouse infected with YK41  
CC with myeloma P3/NS1/1-Ag4-1 to produce a MAb expressing hybridoma.  
CC The DNA obt. was then fused with the expression vector pDA431 to  
CC give PCPN533alpha. The plasmid was used to transform an E. coli  
CC host, which was cultured to give the antigenic polypeptide,  
CC polypeptide A. Polypeptide A and primers and probes derived from  
CC its DNA can be used in assays for the detection of polypeptide A  
CC antibodies and DNA, respectively, useful in the diagnosis of  
CC C. pneumoniae infection.  
XX  
SO Sequence 488 AA:  
Query Match 57.5%; Score 285; DB 17; Length 488;  
Best Local Similarity 100.0%; Pred. No. 2.8e-245;  
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 4 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63
Db 1 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 60
OY 64 ASGDKTSTTETETAPQGVAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 123
Db 61 ASGDKTSTTETETAPQGVAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 120
OY 124 EBASKSMESTLESLSAQAQKEVEAVVAALSGKSSGSAAKLETPPEPKPGVTPRSEVI 183
Db 121 EBASKSMESTLESLSAQAQKEVEAVVAALSGKSSGSAAKLETPPEPKPGVTPRSEVI 180
OY 184 ETGLALAKAIOTLGEATKSALSNASTQAADQTNKLGLEKAIKIDKEREYOEKMAE 243
Db 181 ETGLALAKAIOTLGEATKSALSNASTQAADQTNKLGLEKAIKIDKEREYOEKMAE 240
OY 244 QSKDLEGTMDTVNTVMTAVSAIVTIVISIVAAIFTCGAGLAGLAA 288
Db 241 QSKDLEGTMDTVNTVMTAVSAIVTIVISIVAAIFTCGAGLAGLAA 265

```

RESULT 5  
AAR94586  
ID AAR94586 standard; Protein: 259 AA.  
AC AAR94586;  
XX  
DT 11-NOV-1996 (first entry)  
XX  
DE C. pneumoniae polypeptide antigen (polypeptide A) clone 53-3S.  
XX  
XX  
XX KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;  
KW PCPN533alpha; primer; assay; detection; antibody; diagnosis;  
KW Infection; clone.  
XX  
XX OS Chlamydia pneumoniae.  
XX  
XX PN WO9609320-A1.  
XX  
XX PD 28-MAR-1996.  
XX



960 6 1.2 32 22 AAB45739 Human mature BNP.  
 961 6 1.2 33 14 AAR45762 GLU-BNP. Syntheti  
 962 6 1.2 33 14 AAR35490 Tyr-bBNP. Syntheti  
 963 6 1.2 33 14 AAR36787 prion protein regi  
 964 6 1.2 33 14 AAR36789 prion protein regi  
 965 6 1.2 33 16 AAR80170 MSpI-derived poly  
 966 6 1.2 33 18 AAM17683 prion protein pep  
 967 6 1.2 33 19 AAM70279 Generic peptide se  
 968 6 1.2 33 21 AAB14996 Amino acids 109-14  
 969 6 1.2 33 22 AAB82100 Hamster Prp peptid  
 970 6 1.2 33 22 AAB82109 Murine Prp peptid  
 971 6 1.2 33 22 AAB82115 Murine Prp peptid  
 972 6 1.2 33 22 AAB82119 Murine Prp peptid  
 973 6 1.2 33 22 AAB80992 Wild-type human Pc  
 974 6 1.2 33 22 AAB80993 Mutant human Pcgam  
 975 6 1.2 33 23 AAE15591 Hamster Prp peptid  
 976 6 1.2 33 23 AAE15600 Murine Prp peptid  
 977 6 1.2 33 23 AAE15606 Human Prp (HuPrp)  
 978 6 1.2 33 23 AAE15610 Murine Prp mutant  
 979 6 1.2 34 14 AAR45761 Leader-GLU-BNP. S  
 980 6 1.2 34 16 AAM06201 Spider dragline va  
 981 6 1.2 34 20 AAY33400 N. clavipes spidro  
 982 6 1.2 34 22 ABB30513 Peptide #3164 enco  
 983 6 1.2 34 22 AAU09054 Silk fibre forming  
 984 6 1.2 35 22 AAB66314 P. acnes linoleate  
 985 6 1.2 35 22 AAB66314 MSpI-derived poly  
 986 6 1.2 36 16 AAR80183 Bacillus subtilis  
 987 6 1.2 36 21 AAY56352 Hepatitis C capsid  
 988 6 1.2 36 23 ABB04546 Hepatitis C capsid  
 989 6 1.2 37 19 AAB43445 Scorpion toxin-rel  
 990 6 1.2 38 22 ABB28775 Peptide #1426 enco  
 991 6 1.2 38 22 ABB33962 Peptide #1468 enco  
 992 6 1.2 38 22 ABB19401 Protein #1400 enco  
 993 6 1.2 38 22 AAM54726 Human brain expres  
 994 6 1.2 38 22 AAM67122 Human bone marrow  
 995 6 1.2 38 22 AAM14986 Peptide #1420 enco  
 996 6 1.2 38 22 AAM27419 Peptide #1456 enco  
 997 6 1.2 38 22 AAM02711 Peptide #1393 enco  
 998 6 1.2 38 22 ABB36784 Human peptide enco  
 999 6 1.2 39 19 AAM68413 C-terminal portion  
 1000 6 1.2 39 19 AAM42089 Bst106 synthetic po

## ALIGNMENTS

RESULT 1  
 ID AAB08322 standard; Protein: 496 AA.  
 AC AAB08322;  
 DT 04-DEC-2000 (first entry)  
 DE Amino acid sequence of a Chlamydia pneumoniae protein.  
 DE Chlamydia pneumoniae protein; infection; atherosclerosis;  
 KM coronary atherosclerosis.  
 XX Chlamydia pneumoniae.  
 OS  
 PN WO20046359-A2.  
 PD 10-AUG-2000.  
 PF 28-JAN-2000; 2000WO-GB00237.  
 PR 05-FEB-1999; 99GB-0002555.  
 XX (NEUT-) NEUTEC PHARMA PLC.  
 PA Burnie JP, Matthews RC;  
 PI  
 XX

DR WPI; 2000-543485/49.  
 DR N-PSDB; AAA63621.  
 PT New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,  
 PT preventing and treating C. pneumoniae infection and atherosclerosis,  
 PT including coronary atherosclerosis -  
 PS Claim 1; Page 30-31; 35pp; English.  
 XX  
 CC The present sequence represents a Chlamydia pneumoniae protein. The  
 CC protein, immunogenic fragments of it, nucleotide sequences encoding  
 CC it, or inhibitor for specific against it are used to manufacturing  
 CC a medicament for the treatment of infection due to C. pneumoniae.  
 CC An antibody specific against the protein can diagnose a C. pneumoniae  
 CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,  
 CC including coronary atherosclerosis, caused by C. pneumoniae can also  
 CC be prevented or treated.  
 CC  
 XX  
 SQ Sequence 496 AA:

Query Match 100.0%; Score 496; DB 21; Length 496;  
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;  
 Matches 496; Conservative 0; Indels 0;

OY 1 DTMMSISSSGPDMKNINSQVLTSTPGVPOODKLSGNETKOIOGTROGKNTMESDAT 60  
 |||  
 DB 1 DTMMSISSSGPDMKNINSQVLTSTPGVPOODKLSGNETKOIOGTROGKNTMESDAT 60  
 OY 61 IGASGSKDXTSTTTKTETAPQCGVAAGKESSESOKAGADTVGSAATTAATATKIAMQ 120  
 |||  
 DB 61 IGASGSKDXTSTTTKTETAPQCGVAAGKESSESOKAGADTVGSAATTAATATKIAMQ 120  
 OY 121 TSTEEKSKMESTLESLOSLSAOMKEVEAVVVAALSGSSGSAKLEPELPKPGVTPRS 180  
 |||  
 DB 121 TSTEEKSKMESTLESLOSLSAOMKEVEAVVVAALSGSSGSAKLEPELPKPGVTPRS 180  
 OY 181 EVELIELALAKAIQTLGEMTKSALSNYASTQADQTNKLGLEKQAIKIDKEEEYQEMK 240  
 |||  
 DB 181 EVELIELALAKAIQTLGEMTKSALSNYASTQADQTNKLGLEKQAIKIDKEEEYQEMK 240  
 OY 241 AAEQSKDLEGMTDVTNVTWIAVSAITVIAITVIAITFTGAGLAGLAGAAGAAAGGA 300  
 |||  
 DB 241 AAEQSKDLEGMTDVTNVTWIAVSAITVIAITVIAITFTGAGLAGLAGAAGAAAGGA 300  
 OY 301 AGAAATTTAATQTTVOAVVOAVKQAVITAVROAITAIAAVVSGIKAFIKTLVKAIARA 360  
 |||  
 DB 301 AGAAATTTAATQTTVOAVVOAVKQAVITAVROAITAIAAVVSGIKAFIKTLVKAIARA 360  
 OY 361 ISKGISKVFAGKTOMIAKNPKLSKYISLTSKMTVGVGVVVAAPALGKIMQOLSEM 420  
 |||  
 DB 361 ISKGISKVFAGKTOMIAKNPKLSKYISLTSKMTVGVGVVVAAPALGKIMQOLSEM 420  
 OY 421 QONVAQFOKEVGVKLOAADMISMTFOWOASKIASKQGESNEMTKATKGAQILKAY 480  
 |||  
 DB 421 QONVAQFOKEVGVKLOAADMISMTFOWOASKIASKQGESNEMTKATKGAQILKAY 480  
 OY 481 AATSGAIAAGAKTNNF 496  
 |||  
 DB 481 AATSGAIAAGAKTNNF 496

RESULT 2  
 ID AAY35449 standard; Protein: 496 AA.  
 AC AAY35449;  
 DT 13-SEP-1999 (first entry)  
 DE Amino acid sequence of a Chlamydia pneumoniae protein.  
 DE Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KW

814	6	1.2	10	22	AAg83884	Arabidopsis thalia
815	6	1.2	10	22	AAg83886	Arabidopsis thalia
816	6	1.2	10	22	AAg83888	Arabidopsis thalia
817	6	1.2	10	22	AAg86362	Saccharomyces cere
818	6	1.2	10	22	AAg86665	Saccharomyces cere
819	6	1.2	10	22	AAg87960	Saccharomyces cere
820	6	1.2	10	22	AAg87981	Saccharomyces cere
821	6	1.2	10	22	AAg87982	Saccharomyces cere
822	6	1.2	10	22	AAJ03003	Hepatitis C virus
823	6	1.2	10	22	AAJ11581	Human carcinoembry
824	6	1.2	11	20	AAJ90056	Rhodococcus sp. M4
825	6	1.2	12	14	AAJ37636	Sequence of linker
826	6	1.2	12	18	AAJ32561	Prión target regio
827	6	1.2	12	19	AAJ76451	Graminae pollen al
828	6	1.2	12	19	AAJ76452	Graminae pollen al
829	6	1.2	12	20	AAJ88086	Linker 202' used t
830	6	1.2	12	20	AAJ95434	Peptide linker 202
831	6	1.2	12	21	AAJ27666	Linker 202' connec
832	6	1.2	12	21	AAJ07922	Amino acid sequenc
833	6	1.2	12	21	AAJ57241	Sequence of a pept
834	6	1.2	12	21	AAJ80911	Linker 202', for c
835	6	1.2	12	22	AAJ05174	Pain-relieving tar
836	6	1.2	12	22	AAJ62345	Linker peptide SQ
837	6	1.2	12	22	AAJ50849	Fluorescent protei
838	6	1.2	12	23	AAJ014389	Linker peptide for
839	6	1.2	12	23	AAJ76045	Synthetic linker m
840	6	1.2	13	16	AAJ84042	Murine MHC class I
841	6	1.2	14	14	AAJ34643	Mastoparan mutant
842	6	1.2	14	15	AAJ59502	Hydrophilic linker
843	6	1.2	14	15	AAJ59504	Hydrophilic linker
844	6	1.2	14	18	AAJ17680	Prión protein pept
845	6	1.2	14	18	AAJ17680	Prión protein pept
846	6	1.2	14	21	AAJ25400	Paragut binding a
847	6	1.2	14	21	AAJ13360	SOCs box N-termina
848	6	1.2	14	22	AAJ56578	Human SNP related
849	6	1.2	14	22	AAJ96855	Human peptide #130
850	6	1.2	14	22	AAJ98411	Human peptide #168
851	6	1.2	14	22	AAJ05179	Pain-relieving tar
852	6	1.2	14	22	AAJ00613	Human protein frag
853	6	1.2	14	22	AAJ62350	Linker peptide SQ
854	6	1.2	14	22	AAJ68095	Peptide derived fr
855	6	1.2	14	22	AAJ88089	CD66 peptide CD66e
856	6	1.2	14	22	AAJ50854	Fluorescent protei
857	6	1.2	14	22	AAJ55528	Peptide linker #2.
858	6	1.2	14	22	AAJ57375	Flexible polylinke
859	6	1.2	14	22	AAJ57785	Linker peptide for
860	6	1.2	14	23	AAJ014393	Synthetic linker m
861	6	1.2	14	23	AAJ76049	Hydrophilic linker
862	6	1.2	15	15	AAJ95820	HLA class II bindi
863	6	1.2	15	21	AAJ98912	HLA class II bindi
864	6	1.2	15	21	AAJ98912	Human cathepsin 29
865	6	1.2	15	23	AAJ82648	Epitope mapping st
866	6	1.2	15	23	AAJ82648	Synthetic alternat
867	6	1.2	16	19	AAJ43914	Hepatitis GB virus
868	6	1.2	16	19	AAJ09114	HVRI peptide W1 ha
869	6	1.2	17	23	AAJ67112	Alpha-nellix-formin
870	6	1.2	18	10	AAJ90323	B-cell epitopic pe
871	6	1.2	18	11	AAJ07530	Peptide Mc6.1 isol
872	6	1.2	18	19	AAJ67021	Monomordica charant
873	6	1.2	18	19	AAJ33279	Prión protein regi
874	6	1.2	19	14	AAJ36792	Prión protein regi
875	6	1.2	19	14	AAJ17681	Prión protein pept
876	6	1.2	19	18	AAJ17681	B-cell epitopic pe
877	6	1.2	20	11	AAJ07528	Horseshoe crab rec
878	6	1.2	20	22	AAJ60944	Synthetic HAV pept
879	6	1.2	20	22	AAJ69493	Leucine-rich repea
880	6	1.2	20	23	AAJ75276	Peptide monomer 13
881	6	1.2	21	13	AAJ25975	Human brain natrin
882	6	1.2	21	18	AAJ25774	Antimicrobial olig
883	6	1.2	21	21	AAJ21615	Antimicrobial antimicr
884	6	1.2	21	21	AAJ92801	Human PrP 106-126.
885	6	1.2	21	22	AAJ47107	Synthetic HAV pept
886	6	1.2	21	22	AAJ69494	
887	6	1.2	21	22	AAJ66319	P. acnes linoleate
888	6	1.2	22	19	AAJ71153	Antibacterial and
889	6	1.2	22	20	AAJ06705	Ab5 variable heavy
890	6	1.2	23	21	AAJ14995	Amino acids 106-12
891	6	1.2	23	22	AAJ43915	Peptide #11421 enc
892	6	1.2	23	22	AAJ26833	Protein #8832 enc
893	6	1.2	23	22	AAJ64905	Human brain expres
894	6	1.2	23	22	AAJ77641	Human bone marrow
895	6	1.2	23	22	AAJ21563	Peptide #7997 enco
896	6	1.2	23	22	AAJ37857	Peptide #11874 enc
897	6	1.2	23	22	AAJ82099	Hamster PrP peptid
898	6	1.2	23	23	AAJ46673	Hamster PrP peptid
899	6	1.2	23	23	AAJ15590	Hamster PrP peptid
900	6	1.2	24	15	AAJ62550	Hamster PrP peptid
901	6	1.2	24	18	AAJ13525	Antimicrobial pept
902	6	1.2	24	19	AAJ66411	Anti-melanoma 119n
903	6	1.2	24	21	AAJ69438	Cationic peptide b
904	6	1.2	24	21	AAJ69438	Antimicrobial pept
905	6	1.2	24	21	AAJ91710	Antimicrobial pept
906	6	1.2	24	21	AAJ41243	Cationic peptide B
907	6	1.2	24	22	AAJ43633	Peptide #8749 enco
908	6	1.2	24	22	AAJ25242	Peptide #11139 enc
909	6	1.2	24	22	AAJ26585	Protein #7241 enco
910	6	1.2	24	22	AAJ62109	Protein #8584 enco
911	6	1.2	24	22	AAJ64574	Human brain expres
912	6	1.2	24	22	AAJ74912	Human brain expres
913	6	1.2	24	22	AAJ21300	Human bone marrow
914	6	1.2	24	22	AAJ35027	Peptide #7734 enco
915	6	1.2	24	22	AAJ46799	Peptide #9064 enco
916	6	1.2	24	23	AAJ44671	Human brain natrin
917	6	1.2	25	22	AAJ882547	Human peptide enco
918	6	1.2	25	23	AAJ77599	Human brain natrin
919	6	1.2	26	21	AAJ67295	Human cancer assoc
920	6	1.2	26	21	AAJ99364	Human brain natrin
921	6	1.2	27	21	AAJ59119	Proteasome digesti
922	6	1.2	27	21	AAJ59122	N. clavipes spider
923	6	1.2	27	23	AAJ79178	N. clavipes spider
924	6	1.2	28	17	AAJ89928	Human cytochromealov
925	6	1.2	28	17	AAJ79185	A. cellulolyticus
926	6	1.2	29	14	AAJ36788	Human cytochromealov
927	6	1.2	29	14	AAJ14997	Prión protein regi
928	6	1.2	29	21	AAJ02689	Amino acids 113-14
929	6	1.2	29	22	AAJ82101	Novel human diagno
930	6	1.2	29	23	AAJ804545	Hamster PrP peptid
931	6	1.2	29	23	AAJ15592	Hepatitis C capsid
932	6	1.2	30	21	AAJ67297	Hamster PrP peptid
933	6	1.2	30	21	AAJ67297	Human brain natrin
934	6	1.2	30	22	AAJ27143	Peptide #11784 enc
935	6	1.2	30	22	AAJ65312	Protein #9142 enco
936	6	1.2	30	22	AAJ78012	Human brain expres
937	6	1.2	30	22	AAJ21901	Human bone marrow
938	6	1.2	30	22	AAJ38228	Peptide #8335 enco
939	6	1.2	30	23	AAJ47025	Peptide #12265 enc
940	6	1.2	30	23	AAJ084719	Human peptide enco
941	6	1.2	30	23	AAJ084720	HCV HepC1a segment
942	6	1.2	31	14	AAJ34302	HCV HepC1a segment
943	6	1.2	31	14	AAJ39868	Mutated BMP.
944	6	1.2	31	14	AAJ39868	C peptide RV-C1, r
945	6	1.2	31	23	AAJ79177	Human cytochromealov
946	6	1.2	31	23	AAJ79177	Partial peptide #4
947	6	1.2	32	14	AAJ40861	BNP. Homo sapiens
948	6	1.2	32	14	AAJ43401	Mutated hBNP. Hom
949	6	1.2	32	14	AAJ36381	Recombinant hBNP.
950	6	1.2	32	19	AAJ70090	Brain natriuretic
951	6	1.2	32	19	AAJ51305	Brain natriuretic
952	6	1.2	32	19	AAJ51305	Human B-type natr
953	6	1.2	32	21	AAJ80212	Human B-type natr
954	6	1.2	32	21	AAJ80212	Human B-type natr
955	6	1.2	32	22	AAJ12434	Human B-type natr
956	6	1.2	32	22	AAJ82550	Human brain natrin
957	6	1.2	32	22	AAJ91333	Human brain natrin
958	6	1.2	32	22	AAJ91340	Brain natriuretic
959	6	1.2	32	22	AAJ70810	Human natriuretic

668	7	1.4	766	22	AA682742	S. epidermidis ope	741	7	1.4	1369	23	AAE21712	Human PRIN-7 prote	
669	7	1.4	773	23	AAAM50045	N. clavipes spidro	742	7	1.4	1478	22	ABB63704	Drosophila melanog	
670	7	1.4	777	23	ABBS50046	N. clavipes spidro	743	7	1.4	1486	22	ABB54623	Lactococcus lactis	
671	7	1.4	779	23	ABBS5246	Lactococcus lactis	744	7	1.4	1514	23	ABB92289	Herbicidally activ	
672	7	1.4	780	20	AAV05192	S. aureus pria pro	745	7	1.4	1550	22	ABG09297	Novel human diagno	
673	7	1.4	784	22	ABG15469	Novel human diagno	746	7	1.4	1557	22	ABB62485	Drosophila melanog	
674	7	1.4	788	20	AAV18095	Full length mouse	747	7	1.4	1578	22	ABB71530	Drosophila melanog	
675	7	1.4	793	22	ABBB71688	Drosophila melanog	748	7	1.4	1600	23	ABG66678	Human novel polyipe	
676	7	1.4	796	23	ABP26916	Streptococcus poly	749	7	1.4	1617	23	AAU11794	Dragline protein 1	
677	7	1.4	797	23	ABBA7691	Listeria monocytog	750	7	1.4	1626	23	AAU11798	Dragline protein 1	
678	7	1.4	803	23	ABP26917	Streptococcus poly	751	7	1.4	1640	23	ABBS4727	Lactococcus lactis	
679	7	1.4	804	21	AAV5803	S. pneumoniae pria	752	7	1.4	1852	22	ABB67170	Drosophila melanog	
680	7	1.4	804	22	ABG20282	Novel human diagno	753	7	1.4	1873	22	ABBS59270	Drosophila melanog	
681	7	1.4	805	23	ABP39976	Staphylococcus epi	754	7	1.4	1880	23	AAAM50039	N. clavipes spidro	
682	7	1.4	809	23	AAU11793	Dragline protein 1	755	7	1.4	1911	22	ABM61038	Drosophila melanog	
683	7	1.4	810	21	AAAS5341	Arabidopsis thalia	756	7	1.4	1980	21	AAAB2943	Human GAMP protein	
684	7	1.4	813	22	ABBS59333	Drosophila melanog	757	7	1.4	2006	22	ABB71277	Drosophila melanog	
685	7	1.4	818	22	AAU11797	Dragline protein 1	758	7	1.4	2130	21	AAAS0617	Arabidopsis thalia	
686	7	1.4	823	21	AAAS5340	Arabidopsis thalia	759	7	1.4	2158	21	AAAS0616	Arabidopsis thalia	
687	7	1.4	828	22	ABBB65774	Drosophila melanog	760	7	1.4	2175	22	ABBB65698	Drosophila melanog	
688	7	1.4	831	22	AAAM39264	Human polypeptide	761	7	1.4	2204	21	AAAS0615	Arabidopsis thalia	
689	7	1.4	838	22	AAAM41050	Human polypeptide	762	7	1.4	2240	15	AAAB67819	Acetyl CoA carboxy	
690	7	1.4	844	22	ABBS3083	Escherichia coli p	763	7	1.4	2254	16	AAAB76949	ACCCase. Arabidops	
691	7	1.4	856	22	AAU35717	Helicobacter pylor	764	7	1.4	2254	20	AAV40598	Arabidopsis cytozo	
692	7	1.4	856	22	AAU35896	Helicobacter pylor	765	7	1.4	2254	20	AAV22129	Arabidopsis acetyl	
693	7	1.4	885	22	AAU32393	Novel human secret	766	7	1.4	2257	23	AAAM50332	Alfalfa acetyl-CoA	
694	7	1.4	916	22	ABG18153	Novel human diagno	767	7	1.4	2263	22	ABBB6793	Drosophila melanog	
695	7	1.4	917	21	AAV58637	Protein regulating	768	7	1.4	2290	22	ABBB60066	Drosophila melanog	
696	7	1.4	924	18	AAW00931	Human ICM-4. Hom	769	7	1.4	2293	21	AAAB69197	Human acid sequenc	
697	7	1.4	924	19	AAW60161	Human Intercellula	770	7	1.4	2293	21	AAAB91251	Herbicidally activ	
698	7	1.4	924	19	AAW59006	Human ICM-4 prote	771	7	1.4	2439	22	ABBB68896	Drosophila melanog	
699	7	1.4	924	20	AAV05467	Human ICM-4 prote	772	7	1.4	2675	21	AAAB07564	Protein encoded by	
700	7	1.4	931	20	AAW73519	Human neuronal ICA	773	7	1.4	2809	23	AAAG66169	Human fibrillin 3	
701	7	1.4	934	23	ABBB09319	Human transient re	774	7	1.4	2837	22	ABBS7791	Drosophila melanog	
702	7	1.4	938	22	ABG10442	Novel human diagno	775	7	1.4	2958	22	ABB71368	Drosophila melanog	
703	7	1.4	943	12	AAAI3802	170kD heavy chain	776	7	1.4	2972	21	AAAB09265	Hepatitis GB virus	
704	7	1.4	950	22	AAOI2744	Human polypeptide	777	7	1.4	3163	16	AAAB94347	Beta-amyloid pepti	
705	7	1.4	963	22	AAU36226	Pseudomonas aerugi	778	7	1.4	3503	22	ABBB6499	Drosophila melanog	
706	7	1.4	977	22	AAAB82315	Human immunoglobul	779	7	1.4	3680	22	ABG22108	Novel human diagno	
707	7	1.4	988	22	ABBB60007	Drosophila melanog	780	7	1.4	4551	21	AAAB18637	Amino acid sequenc	
708	7	1.4	989	23	AAAM50038	N. clavipes spidro	781	7	1.4	4551	21	AAAB18637	Narboxilide synth	
709	7	1.4	992	16	AAAB73917	Ribellia virus Ther	782	7	1.4	4613	21	AAV77192	S. venezuelae macr	
710	7	1.4	994	19	AAW44844	S. lepidophylla tr	783	7	1.4	4613	21	AAV77200	S. venezuelae pik	
711	7	1.4	997	22	ABG09929	Novel human diagno	784	7	1.4	12199	21	AAV77180	Human titin (conn	
712	7	1.4	1007	22	ABBS7807	Drosophila melanog	785	7	1.4	26926	22	AAU05396	Human titin (conn	
713	7	1.4	1012	23	ABBB90839	Herbicidally activ	786	7	1.2		21	AAAB08327	Epitope derived fr	
714	7	1.4	1013	22	AAU32392	Novel human secret	787	7	1.2		6	21	AAAB08329	Epitope derived fr
715	7	1.4	1029	23	AAU93080	Arabidopsis transc	788	7	1.2		8	16	AAAB78841	P. falciparum glrc
716	7	1.4	1035	22	ABBB63917	Drosophila melanog	789	7	1.2		8	20	AAAB89379	Beta-amyloid pepti
717	7	1.4	1062	22	ABBB71830	Drosophila melanog	790	7	1.2		8	22	AAAB82093	Hamster Prp peptid
718	7	1.4	1095	22	ABG21602	Novel human diagno	791	7	1.2		6	22	ABBB05165	Prion protein pept
719	7	1.4	1095	22	ABG21823	Novel human diagno	792	7	1.2		8	23	AAE15584	Hamster Prp peptid
720	7	1.4	1118	22	AAAM50209	Human fibrillin-1i	793	7	1.2		9	18	AAW43905	Specific human Ieu
721	7	1.4	1122	21	AAV32218	Drosophila PER B p	794	7	1.2		9	18	AAW43918	Specific human Ieu
722	7	1.4	1127	22	AAU68562	Human novel cytoki	795	7	1.2		9	19	AAW53354	Nephila clavipes s
723	7	1.4	1155	21	AAV32219	Drosophila PER C p	796	7	1.2		9	19	AAW47782	Peptide prepared u
724	7	1.4	1193	23	ABBB90860	Herbicidally activ	797	7	1.2		9	20	AAV48090	Immunogenic peptid
725	7	1.4	1201	22	ABG10512	Novel human diagno	798	7	1.2		9	21	AAAB08331	Epitope derived fr
726	7	1.4	1218	21	AAV32217	Drosophila PER A p	799	7	1.2		9	21	AAV59079	Second segment of
727	7	1.4	1218	22	ABBS59285	Drosophila melanog	800	7	1.2		10	15	AAV38019	Influenza virus-de
728	7	1.4	1232	22	ABBS58856	Drosophila melanog	801	7	1.2		10	17	AAAR92136	Porphyromonas ging
729	7	1.4	1235	17	AAW21572	Rat brain serine e	802	7	1.2		10	20	AAV45589	Immunogenic peptid
730	7	1.4	1237	22	ABG21871	Novel human diagno	803	7	1.2		10	22	ABP12785	HIV A02 super moti
731	7	1.4	1249	17	AAW21571	Rat brain homologo	804	7	1.2		10	22	ABP17791	HIV B58 super moti
732	7	1.4	1255	23	AAAM50037	N. clavipes spidro	805	7	1.2		10	22	AAAG94912	Human complementar
733	7	1.4	1276	22	AAAB70234	Novel human diagno	806	7	1.2		10	22	AAAG94912	Human complementar
734	7	1.4	1277	22	AAAB70236	Novel human diagno	807	7	1.2		10	22	AAAG97195	Human complementar
735	7	1.4	1280	22	ABG19587	Entamoeba histolyt	808	7	1.2		10	22	AAAG97197	Human complementar
736	7	1.4	1291	16	AAAB68975	Entamoeba histolyt	809	7	1.2		10	22	AAAG97199	Human complementar
737	7	1.4	1291	22	AAAB70233	Entamoeba histolyt	810	7	1.2		10	22	AAAG97712	Human complementar
738	7	1.4	1292	16	AAAB70681	Entamoeba histolyt	811	7	1.2		10	22	AAAG97945	Human complementar
739	7	1.4	1292	22	AAAB70235	Salmonella typhl c	812	7	1.2		10	22	AAAG97946	Human complementar
740	7	1.4	1362	22	AAU38416	Salmonella typhl c	813	7	1.2		10	22	AAAG97948	Human complementar

522	7	1.4	442	21	AAV97176	48 kDa immobilizat	595	7	1.4	559	22	AAU08660	Human NOV4 protein
523	7	1.4	443	14	AAAR43895	heL-nl. Homo sapi	596	7	1.4	563	22	ABR63647	Drosophila melanog
524	7	1.4	443	20	AAAY05193	S. aureus pria pro	597	7	1.4	563	21	AAAG31276	Arabidopsis thalia
525	7	1.4	444	19	AAAM60579	Drosophila neuron-	598	7	1.4	566	22	AAU27967	Human contig polyP
526	7	1.4	445	20	AAAG67830	Human muscarinic a	599	7	1.4	566	22	AAU27968	Human contig polyP
527	7	1.4	445	20	AAAG67831	Rat muscarinic ace	600	7	1.4	566	23	AAAB92932	Hereticidally activ
528	7	1.4	445	20	AAAY06322	Human g protein co	601	7	1.4	568	22	AAAB66501	Drosophila melanog
529	7	1.4	445	20	AAAY06323	Rat G protein comp	602	7	1.4	572	19	AAAM66391	Aspergillus fumiga
530	7	1.4	445	20	AAAY92975	Human MACHR-6 prot	603	7	1.4	572	21	AAAG31517	Arabidopsis thalia
531	7	1.4	445	21	AAAB15382	Rat G-protein coup	604	7	1.4	573	21	AAAG42355	Arabidopsis thalia
532	7	1.4	445	21	AAAY92218	Human histamine H3	605	7	1.4	578	21	AAAY15187	Rice phytoene desa
533	7	1.4	445	22	AAAG30627	A human histamine	606	7	1.4	578	22	AAAB59076	Drosophila melanog
534	7	1.4	448	22	AAAG90237	C glutamic prote	607	7	1.4	580	23	AAAG00330	Novel human diagno
535	7	1.4	448	22	AAAG90046	C glutamic prote	608	7	1.4	590	23	AAAB29967	Streptococcus poly
536	7	1.4	449	22	AAAB80082	Corynebacterium gl	609	7	1.4	592	23	AAAB29967	Streptococcus poly
537	7	1.4	449	22	ABG12680	Novel human diagno	610	7	1.4	598	19	AAAM46313	AAV4 VP2 coat prot
538	7	1.4	451	23	AAAB60688	Phaseolin promoter	611	7	1.4	598	19	AAAM46313	AAV4 VP2 coat prot
539	7	1.4	453	23	AAAY22208	Human G-protein co	612	7	1.4	600	19	AAAM39080	S. lepidophylla tr
540	7	1.4	453	21	AAAB15381	Human G-protein co	613	7	1.4	604	16	AAAR99057	Spider dragline va
541	7	1.4	457	19	AAAM42430	Escherichia coli g	614	7	1.4	604	23	AAAB25463	Spider dragline va
542	7	1.4	457	19	AAAB60694	Mycobacterium lepr	615	7	1.4	606	16	AAAR99053	Spider dragline va
543	7	1.4	458	23	AAAB60702	Mycobacterium lepr	616	7	1.4	606	16	AAAR99055	Spider dragline va
544	7	1.4	458	23	AAAB60778	Mycobacterium lepr	617	7	1.4	606	20	AAAY40100	Polymer of an anal
545	7	1.4	458	23	AAAB60865	Mycobacterium lepr	618	7	1.4	606	20	AAAY40101	Polymer of an anal
546	7	1.4	459	21	AAAG24771	Arabidopsis thalia	619	7	1.4	606	20	AAAY40102	Polymer of an anal
547	7	1.4	459	21	AAAG50620	Arabidopsis thalia	620	7	1.4	615	22	AAAU51381	Propionibacterium
548	7	1.4	460	21	AAAG25456	Arabidopsis thalia	621	7	1.4	617	22	AAAB65168	Drosophila melanog
549	7	1.4	460	21	AAAG48339	Arabidopsis thalia	622	7	1.4	617	23	AAAB92791	Hereticidally activ
550	7	1.4	466	21	AAAG24770	Arabidopsis thalia	623	7	1.4	623	22	AAAB85755	A. thaliana rd22BP
551	7	1.4	466	21	AAAG50619	Arabidopsis thalia	624	7	1.4	623	22	AAAB85760	Hereticidally activ
552	7	1.4	471	23	AAAB60695	Arabidopsis thalia	625	7	1.4	623	23	AAAB91226	Hereticidally activ
553	7	1.4	472	22	AAAB69894	Drosophila melanog	626	7	1.4	625	21	AAAY97397	Mouse apoptosis-in
554	7	1.4	476	22	AAU34421	E. coli cellular p	627	7	1.4	625	22	AAAB58966	Drosophila melanog
555	7	1.4	478	22	AAAB69816	Drosophila melanog	628	7	1.4	626	18	AAAM16398	Human neuron-deriv
556	7	1.4	481	20	AAAY50342	Murine Kv6.2 prote	629	7	1.4	630	23	AAAM50042	N. clavipes spidro
557	7	1.4	481	21	AAAG25455	Arabidopsis thalia	630	7	1.4	631	21	AAAG29253	Arabidopsis thalia
558	7	1.4	481	21	AAAG48338	Arabidopsis thalia	631	7	1.4	632	21	AAAB41722	Human ORFX ORF1486
559	7	1.4	482	22	AAAB50165	Human transcriptio	632	7	1.4	632	22	AAAB93818	Human polypeptide.
560	7	1.4	482	22	AAAB70692	Human DEC2a protei	633	7	1.4	632	22	AAAB65651	Novel protein kina
561	7	1.4	483	20	AAAY31656	Mouse N-acetylgluc	634	7	1.4	634	22	AAAB61272	Drosophila melanog
562	7	1.4	484	22	AAAM41276	Human polypeptide	635	7	1.4	634	22	AAAB66397	Drosophila melanog
563	7	1.4	484	22	AAAM41277	Human polypeptide	636	7	1.4	634	22	AAAB66396	Drosophila melanog
564	7	1.4	489	22	AAAB70693	Human DEC2b protei	637	7	1.4	636	22	AAAG03757	Novel human diagno
565	7	1.4	489	22	AAAB64492	Drosophila melanog	638	7	1.4	640	22	AAAG29028	Novel human diagno
566	7	1.4	502	22	AAAB71952	Drosophila melanog	639	7	1.4	643	20	AAAY14964	Amino acid sequenc
567	7	1.4	506	22	AAAB71952	Novel human diagno	640	7	1.4	643	22	AAAB83264	C heterostrophus F
568	7	1.4	512	19	AAAM68473	HIV-1 strain YBF30	641	7	1.4	648	22	AAAB19586	Novel human diagno
569	7	1.4	513	23	AAAB60691	Phaseolin promoter	642	7	1.4	655	23	AAAB76672	Protein related to
570	7	1.4	517	23	AAAB41178	Human ovarian anti	643	7	1.4	660	22	AAAB62471	Drosophila melanog
571	7	1.4	518	23	AAAB60693	Promoter-Oleoin t	644	7	1.4	668	22	AAAB66214	Drosophila melanog
572	7	1.4	519	22	AAAB68337	Drosophila melanog	645	7	1.4	668	22	AAAG30062	Novel human diagno
573	7	1.4	520	13	AAAR21075	Sequence encoded b	646	7	1.4	673	21	AAAG43844	Arabidopsis thalia
574	7	1.4	520	13	AAAR21076	Sequence encoded b	647	7	1.4	673	23	AAAB59674	Drosophila melanog
575	7	1.4	520	13	AAAR21077	Sequence encoded b	648	7	1.4	676	23	AAAM50047	N. clavipes spidro
576	7	1.4	520	22	AAU41718	Propionibacterium	649	7	1.4	677	22	AAAB61500	Drosophila melanog
577	7	1.4	521	21	AAAG31518	Arabidopsis thalia	650	7	1.4	681	22	AAAB82609	Spider recombinant
578	7	1.4	526	22	AAAB99895	Xanthomons sp. ch	651	7	1.4	686	21	AAAG43843	Arabidopsis thalia
579	7	1.4	528	22	AAAM78918	Human protein SEO	652	7	1.4	689	22	AAAB71516	Drosophila melanog
580	7	1.4	528	22	AAAB82611	Spider recombinant	653	7	1.4	691	22	AAAB82610	Spider recombinant
581	7	1.4	529	21	AAAY44989	Human epidermal pr	654	7	1.4	702	23	AAAM47575	Drosophila cell cy
582	7	1.4	530	21	AAAG24769	Arabidopsis thalia	655	7	1.4	709	13	AAAR24130	Toxic KHS gene. S
583	7	1.4	535	22	AAAG29491	Novel human diagno	656	7	1.4	714	22	AAAB12681	Novel human diagno
584	7	1.4	537	21	AAAG31277	Arabidopsis thalia	657	7	1.4	718	22	AAAB19588	Novel human diagno
585	7	1.4	537	22	AAAB12238	Human secreted pro	658	7	1.4	724	21	AAAG43842	Arabidopsis thalia
586	7	1.4	542	22	AAAB84706	Nucleotide sequenc	659	7	1.4	725	22	AAAM78889	Human protein SEO
587	7	1.4	543	22	AAAB12660	Novel human diagno	660	7	1.4	732	22	AAAM79873	Human protein SEO
588	7	1.4	544	19	AAAM46314	AAV4 VP3 coat prot	661	7	1.4	734	19	AAAM46308	AAV4 VP1 capsid pr
589	7	1.4	546	21	AAAG50618	Arabidopsis thalia	662	7	1.4	738	19	AAAM56163	New DNA sequence i
590	7	1.4	547	22	AAAB63123	Drosophila melanog	663	7	1.4	739	22	AAAB68512	Drosophila melanog
591	7	1.4	550	22	AAAM41994	Human polypeptide	664	7	1.4	739	22	AAAB11417	Human secreted pro
592	7	1.4	550	22	AAAY97630	Rice geranylgerany	665	7	1.4	751	18	AAAM13491	Helicobacter ClpB
593	7	1.4	553	22	AAAB62266	Drosophila melanog	666	7	1.4	758	20	AAAY35578	C. pneumoniae prot
594	7	1.4	556	23	AAAB48441	Listeria monocytog	667	7	1.4	762	20	AAAY37364	Protein involved i

376	7	1.4	258	22	ABG17521	Novel human diagno	449	7	1.4	333	23	ABP60685	Phaseolin promoter
377	7	1.4	260	21	AAG47476	Arabidopsis thalia	450	7	1.4	333	23	ABP60846	Arabidopsis thalia
378	7	1.4	260	22	ABB67475	Drosophila melanog	451	7	1.4	333	23	ABB93254	Herbicidally activ
379	7	1.4	261	14	AAR44230	Chimeric Ig superf	452	7	1.4	333	23	AAE18732	Arabidopsis NADPH
380	7	1.4	263	12	AAR11185	Plasmid pBTa870-en	453	7	1.4	333	23	ABB04150	Recombinant Arabid
381	7	1.4	264	23	AAM50048	N. clavipes spidro	454	7	1.4	335	22	ABG23968	Novel human diagno
382	7	1.4	266	22	AAU35242	Enterococcus faeca	455	7	1.4	340	22	AAB82911	Human histamine H3
383	7	1.4	267	22	ABB64636	Drosophila melanog	456	7	1.4	341	22	AAE05465	Human Par-4 protei
384	7	1.4	271	22	AAG91903	C glutamicum prote	457	7	1.4	344	21	AAV74703	Nisseria meningit
385	7	1.4	271	23	AAM50044	N. clavipes spidro	458	7	1.4	344	22	ABG06973	Novel human diagno
386	7	1.4	271	23	AAU93059	Arabidopsis transc	459	7	1.4	344	22	AAB79891	Corynebacterium g1
387	7	1.4	272	21	AAG12672	Arabidopsis thalia	460	7	1.4	345	21	AAV16955	Arabidopsis thalia
388	7	1.4	274	21	AAG16956	Arabidopsis thalia	461	7	1.4	345	21	AAV74701	Neisseria gonorrhe
389	7	1.4	278	22	AAU17486	Novel signal trans	462	7	1.4	350	22	AAB65881	Drosophila melanog
390	7	1.4	279	21	AAG16375	Arabidopsis thalia	463	7	1.4	351	22	AAU35596	Haemophilus influe
391	7	1.4	279	21	AAG21798	Arabidopsis thalia	464	7	1.4	351	22	AAG65581	Human histamine H3
392	7	1.4	282	22	AAB47283	MST-3793. Mycobac	465	7	1.4	351	22	AAU99817	Human histamine H3
393	7	1.4	283	12	AAR11186	Plasmid pBTa862-en	466	7	1.4	353	23	ABG05551	Novel human diagno
394	7	1.4	285	22	AAU50342	Protonibacterium	467	7	1.4	353	22	ABG06757	Novel human diagno
395	7	1.4	287	16	AAR74201	AUF1 polypeptide.	468	7	1.4	362	21	AAB15769	Nicotiana plumbagi
396	7	1.4	289	21	AAG20045	Arabidopsis thalia	469	7	1.4	364	23	AAM50043	N. clavipes spidro
397	7	1.4	291	22	ABB58596	Drosophila melanog	470	7	1.4	365	21	AAG36682	Arabidopsis thalia
398	7	1.4	291	22	AAU35321	Enterococcus faeca	471	7	1.4	365	22	AAG65580	Human histamine H3
399	7	1.4	291	22	AAB82608	Spider recombinant	472	7	1.4	365	23	ABR92366	Herbicidally activ
400	7	1.4	293	21	AAG14282	Arabidopsis thalia	473	7	1.4	366	21	AAV54993	Arabidopsis thalia
401	7	1.4	293	21	AAG20267	Arabidopsis thalia	474	7	1.4	368	21	AAG21064	Arabidopsis thalia
402	7	1.4	294	22	ABG17009	Novel human diagno	475	7	1.4	369	21	AAG20265	Arabidopsis thalia
403	7	1.4	299	22	ABB70132	Drosophila melanog	476	7	1.4	374	20	AAV43981	Horse alcohol dehy
404	7	1.4	299	22	AAU30809	Novel human secret	477	7	1.4	374	20	AAV43982	Horse alcohol dehy
405	7	1.4	300	21	AAG36684	Arabidopsis thalia	478	7	1.4	374	20	AAV43983	Human alcohol dehy
406	7	1.4	301	22	AAU00612	F. venenatum modif	479	7	1.4	374	20	AAV43984	Human alcohol dehy
407	7	1.4	301	22	AAU00959	F. venenatum glucu	480	7	1.4	374	20	AAV43985	Human alcohol dehy
408	7	1.4	301	22	AAU00960	F. sporotrichioid	481	7	1.4	374	20	AAV43986	Mouse alcohol dehy
409	7	1.4	301	22	AAU02018	F. venenatum modif	482	7	1.4	374	22	AAU28328	Novel human secret
410	7	1.4	303	22	ABB67154	Drosophila melanog	483	7	1.4	375	21	AAG52946	Arabidopsis thalia
411	7	1.4	305	21	AAG05380	Arabidopsis thalia	484	7	1.4	376	19	AAW81744	M. tuberculosis im
412	7	1.4	305	21	AAG36683	Arabidopsis thalia	485	7	1.4	376	19	AAW64377	Mycobacterium tube
413	7	1.4	305	21	AAG48337	Arabidopsis thalia	486	7	1.4	376	20	AAV39174	M. tuberculosis an
414	7	1.4	305	21	AAV58038	Rice chitinase pro	487	7	1.4	376	20	AAV39031	M. tuberculosis re
415	7	1.4	306	22	ABB58203	Drosophila melanog	488	7	1.4	378	21	AAG32195	Arabidopsis thalia
416	7	1.4	310	21	AAG47475	Arabidopsis thalia	489	7	1.4	378	21	AAG45167	Arabidopsis thalia
417	7	1.4	310	22	ABR71086	Drosophila melanog	490	7	1.4	379	21	AAG45166	Arabidopsis thalia
418	7	1.4	310	23	AAE18733	Rice NADPH depende	491	7	1.4	383	21	AAG32194	Arabidopsis thalia
419	7	1.4	311	23	ABP60867	Mycobacterium smeg	492	7	1.4	383	23	ABP60847	Arabidopsis thalia
420	7	1.4	313	21	AAB14138	Bordetella pertuss	493	7	1.4	385	21	AAB53426	Human colon cancer
421	7	1.4	318	21	AAG48300	Arabidopsis thalia	494	7	1.4	385	22	ABG73824	Human colon cancer
422	7	1.4	318	22	ABB64755	Drosophila melanog	495	7	1.4	388	22	AAB96706	putative P. abyssi
423	7	1.4	320	22	AAM00849	Human bone marrow	496	7	1.4	388	23	ABP60687	Phaseolin promoter
424	7	1.4	321	22	ABB66897	Drosophila melanog	497	7	1.4	391	21	AAG42557	Arabidopsis thalia
425	7	1.4	321	23	ABP60873	Streptomyces clavu	498	7	1.4	391	22	AAM39491	Human polypeptide
426	7	1.4	321	23	ABP60874	Streptomyces coeli	499	7	1.4	395	21	AAG25457	Arabidopsis thalia
427	7	1.4	322	21	AAG07651	Arabidopsis thalia	500	7	1.4	395	21	AAG48340	Arabidopsis thalia
428	7	1.4	322	21	AAG39412	Arabidopsis thalia	501	7	1.4	398	21	AAG42556	Arabidopsis thalia
429	7	1.4	323	12	AAR11187	Plasmid pBTa859-en	502	7	1.4	401	23	ABB49278	Listeria monocytop
430	7	1.4	323	22	AAM93382	Human polypeptide,	503	7	1.4	402	8	AAV70709	Plasmodium cynomol
431	7	1.4	325	21	AAG07650	Arabidopsis thalia	504	7	1.4	403	22	ABB58558	Drosophila melanog
432	7	1.4	325	21	AAG39411	Arabidopsis thalia	505	7	1.4	403	22	AAG52945	Arabidopsis thalia
433	7	1.4	326	23	ABP60917	Streptomyces coeli	506	7	1.4	408	22	ABB68464	Drosophila melanog
434	7	1.4	330	21	AAG32196	Arabidopsis thalia	507	7	1.4	409	21	AAB25889	IM48 (G1) surfact
435	7	1.4	330	21	AAG39410	Arabidopsis thalia	508	7	1.4	413	20	AAV22207	G-protein conjugat
436	7	1.4	330	21	AAG45168	Arabidopsis thalia	509	7	1.4	413	21	AAB15383	Rat G-protein coup
437	7	1.4	331	21	AAG07649	Arabidopsis thalia	510	7	1.4	417	22	ABB61903	Drosophila melanog
438	7	1.4	331	22	AAE10456	E. coli flig prote	511	7	1.4	417	22	ABB66270	Drosophila melanog
439	7	1.4	332	21	AAB29250	Barley NADPH-thior	512	7	1.4	420	21	AAG22299	Arabidopsis thalia
440	7	1.4	332	21	AAB29256	A. thaliana NADPH-	513	7	1.4	420	21	AAG36821	Arabidopsis thalia
441	7	1.4	332	22	ABB62523	Drosophila melanog	514	7	1.4	425	22	AAW78474	Human protein SBO
442	7	1.4	332	23	ABP60678	Arabidopsis NADPH	515	7	1.4	429	22	AAU59019	Protonibacterium
443	7	1.4	332	23	ABP60697	Arabidopsis thalia	516	7	1.4	430	22	AAW79458	Human protein SEO
444	7	1.4	333	21	AAG14281	Arabidopsis thalia	517	7	1.4	430	22	AAM00962	Human bone marrow
445	7	1.4	333	21	AAAG20266	Arabidopsis thalia	518	7	1.4	433	21	AAG22298	Arabidopsis thalia
446	7	1.4	333	21	AAG52947	Novel human diagno	519	7	1.4	437	22	ABG26124	Novel human diagno
447	7	1.4	333	23	ABP60677	Arabidopsis NADPH	520	7	1.4	438	21	AAG36820	Arabidopsis thalia
448	7	1.4	333	23	ABP60679	Arabidopsis NADPH	521	7	1.4	442	21	AAB25859	48KD I-antigen pro

230	7	1.4	103	22	AAU43490	Proprionibacterium
231	7	1.4	104	22	AAH50034	N. clavipes spidro
232	7	1.4	105	22	AAG89907	C glutamicum prote
233	7	1.4	107	21	AAV94312	Soybean PCNA P120
234	7	1.4	107	23	ABB76673	Protein related to
235	7	1.4	109	17	AAH84086	T-lymphocyte stimu
236	7	1.4	109	17	ABG28813	Novel human diagno
237	7	1.4	110	22	AAH99611	Cladosporium herba
238	7	1.4	110	21	AAH44063	Human cancer assoc
239	7	1.4	110	21	AAH30492	Arbidiopsis thalia
240	7	1.4	111	18	AAW22490	Phaffia derived gl
241	7	1.4	111	19	AAW61478	A. fumigatus aller
242	7	1.4	113	15	AAH63386	Hepatitis C virus
243	7	1.4	113	15	AAH55294	HCV isolate BE93 N
244	7	1.4	113	16	AAH71833	Enterococcus faec
245	7	1.4	113	21	AAH16394	Arbidiopsis thalia
246	7	1.4	113	21	AAH19786	Arbidiopsis thalia
247	7	1.4	115	23	ABP10725	Human OREX protei
248	7	1.4	118	21	AAH06255	Arbidiopsis thalia
249	7	1.4	119	21	AAH24148	Arbidiopsis thalia
250	7	1.4	120	23	ABH08473	protein #5 related
251	7	1.4	120	23	ABH49549	Listeria monocytoc
252	7	1.4	124	22	AAH92215	C glutamicum prote
253	7	1.4	125	21	AAH42643	Human OREX protei
254	7	1.4	126	22	AAH77149	Human colon cancer
255	7	1.4	126	21	ABP28369	Streptococcus poly
256	7	1.4	127	21	AAH12649	Arbidiopsis thalia
257	7	1.4	127	21	AAH26891	Zea mays protein f
258	7	1.4	127	21	AAH43378	Exemplary truncate
259	7	1.4	129	21	AAH46659	Arbidiopsis thalia
260	7	1.4	132	17	AAH05211	Canola acetyl-CoA
261	7	1.4	132	19	AAH70408	Canola acetyl-CoA
262	7	1.4	133	22	AAH90929	Cenarchaeum symbio
263	7	1.4	133	22	AAH02545	Human polypeptide
264	7	1.4	134	21	AAH44247	Arbidiopsis thalia
265	7	1.4	137	22	ABH67892	Drosophila melanog
266	7	1.4	137	22	AAH50036	N. clavipes spidro
267	7	1.4	138	21	AAH26890	Zea mays protein f
268	7	1.4	138	22	ABH06660	Novel human diagno
269	7	1.4	141	22	AAH43874	Proprionibacterium
270	7	1.4	142	21	AAH93408	Amino acid sequenc
271	7	1.4	145	22	ABH71046	Drosophila melanog
272	7	1.4	145	22	AAH06189	Streptomyces venez
273	7	1.4	146	18	AAH16312	Streptomyces venez
274	7	1.4	146	21	AAH46658	Arbidiopsis thalia
275	7	1.4	147	19	AAH60173	Heavy chain variab
276	7	1.4	148	21	AAH26134	Adenovirus 5 E2 an
277	7	1.4	149	22	AAH93494	Human polypeptide
278	7	1.4	150	22	ABH68625	Drosophila melanog
279	7	1.4	150	22	AAH44099	Peptide sequence f
280	7	1.4	151	21	AAH44246	Arbidiopsis thalia
281	7	1.4	152	21	AAH16627	Arbidiopsis thalia
282	7	1.4	152	21	AAH20115	Arbidiopsis thalia
283	7	1.4	152	21	AAH27422	Arbidiopsis thalia
284	7	1.4	152	22	ABH69237	Arbidiopsis thalia
285	7	1.4	154	22	AAH12673	Drosophila melanog
286	7	1.4	155	22	AAH41768	Arbidiopsis thalia
287	7	1.4	156	22	AAH27421	Human polypeptide
288	7	1.4	156	22	AAH47181	Arbidiopsis thalia
289	7	1.4	156	22	AAH67407	Proprionibacterium
290	7	1.4	158	21	AAH06254	Arbidiopsis thalia
291	7	1.4	158	21	AAH20046	Arbidiopsis thalia
292	7	1.4	161	21	AAH20114	Arbidiopsis thalia
293	7	1.4	163	21	AAH30169	Arbidiopsis thalia
294	7	1.4	163	21	AAH49016	Arbidiopsis thalia
295	7	1.4	163	21	AAH49941	Arbidiopsis thalia
296	7	1.4	163	23	ABP51038	Human MDR1 SEQ ID
297	7	1.4	164	21	AAH21031	Human nucleic acid
298	7	1.4	164	22	AAH39982	Human polypeptide
299	7	1.4	164	22	AAH33367	Human protein sequ
300	7	1.4	166	21	AAH30975	Arbidiopsis thalia
301	7	1.4	166	21	AAH26889	Arbidiopsis thalia
302	7	1.4	166	21	AAH52922	Arbidiopsis thalia



84	8	1.6	224	22	AAB20105	Moraxella catarrha	157	7	1.4	25	13	AAR22910	TraT(3). Escherich
85	8	1.6	240	21	AAG07289	Arabidopsis thalia	158	7	1.4	25	21	AAV59101	N. clavipes spider
86	8	1.6	243	21	AAG40498	Arabidopsis thalia	159	7	1.4	27	21	AAV59110	N. clavipes spider
87	8	1.6	247	21	AAG38336	Arabidopsis thalia	160	7	1.4	27	21	AAV59114	N. clavipes spider
88	8	1.6	251	22	ABG03796	Novel human diagno	161	7	1.4	27	21	AAV59118	N. clavipes spider
89	8	1.6	254	22	AAB94600	Human protein sequ	162	7	1.4	28	21	AAV59103	N. clavipes spider
90	8	1.6	262	22	AAU23380	Novel human enzyme	163	7	1.4	28	21	AAV59106	N. clavipes spider
91	8	1.6	262	22	AAU21782	Novel human neopla	164	7	1.4	30	21	AAV59104	N. clavipes spider
92	8	1.6	264	21	AAV96273	Rice RA8 anther-sp	165	7	1.4	30	21	AAV59111	N. clavipes spider
93	8	1.6	266	22	AAV90461	C. glutamicum prote	166	7	1.4	30	21	AAV59116	N. clavipes spider
94	8	1.6	300	23	ABB08606	TAP protein SEQ ID	167	7	1.4	30	21	AAV59121	N. clavipes spider
95	8	1.6	327	21	AAG31744	Arabidopsis thalia	168	7	1.4	30	21	AAV59123	N. clavipes spider
96	8	1.6	331	17	AAR98409	Rheumatoid arthrit	169	7	1.4	31	21	AAV59109	N. clavipes spider
97	8	1.6	347	21	AAG07288	Arabidopsis thalia	170	7	1.4	31	22	ABB38864	Peptide #6370 enco
98	8	1.6	350	21	AAG40497	Arabidopsis thalia	171	7	1.4	31	22	ABB23868	Protein #5867 enco
99	8	1.6	354	21	AAG38335	Arabidopsis thalia	172	7	1.4	31	22	AAV59511	Human brain expres
100	8	1.6	367	21	AAV55639	M. tuberculosis yc	173	7	1.4	31	22	AAV72075	Human bone marrow
101	8	1.6	373	21	AAG07287	Arabidopsis thalia	174	7	1.4	31	22	AAV32339	Peptide #6376 enco
102	8	1.6	380	21	AAG38334	Arabidopsis thalia	175	7	1.4	31	23	ABG41891	Human peptide enco
103	8	1.6	407	19	AAV98453	H. pylori GHPD 639	176	7	1.4	32	21	AAV59107	N. clavipes spider
104	8	1.6	408	18	AAV19760	Pyruvate:ferredoxi	177	7	1.4	33	19	AAV53350	N. clavipes spider
105	8	1.6	431	21	AAV19410	An aromatic class	178	7	1.4	33	21	AAV59074	N. clavipes spider
106	8	1.6	431	23	ABB81703	Arabidopsis sp. AT	179	7	1.4	33	22	ABB37838	Peptide #5344 enco
107	8	1.6	431	23	AAU72773	Arabidopsis locoph	180	7	1.4	33	22	ABB23102	Protein #5101 enco
108	8	1.6	434	21	AAG31743	Arabidopsis thalia	181	7	1.4	33	22	AAV58460	Human brain expres
109	8	1.6	469	21	AAG31742	Arabidopsis thalia	182	7	1.4	33	22	AAV70953	Human bone marrow
110	8	1.6	477	22	ABB59816	Drosophila melanog	183	7	1.4	33	22	AAV18744	Peptide #5178 enco
111	8	1.6	505	22	ABG03717	Novel human diagno	184	7	1.4	33	22	AAV31240	Peptide #5277 enco
112	8	1.6	512	23	ABB48169	Listeria monocytos	185	7	1.4	33	23	ABG40751	Human peptide enco
113	8	1.6	549	20	AAV06220	EPEC E. coli trans	186	7	1.4	33	23	AAE18322	Spider silk proteol
114	8	1.6	553	23	ABB59793	Drosophila melanog	187	7	1.4	34	21	AAV59105	N. clavipes spider
115	8	1.6	587	23	ABB54271	Lactococcus lactis	188	7	1.4	34	21	AAV59115	N. clavipes spider
116	8	1.6	632	21	AAV42782	Human ORFX ORF2546	189	7	1.4	34	21	AAV59117	N. clavipes spider
117	8	1.6	640	22	ABB67963	Drosophila melanog	190	7	1.4	47	22	AAV68568	Human novel cytoxi
118	8	1.6	646	18	AAV27178	Nephila clavipes s	191	7	1.4	52	22	AAV48244	Propionibacterium
119	8	1.6	651	20	AAV40097	Spider silk proteol	192	7	1.4	66	21	AAV01698	Human secreted pro
120	8	1.6	651	23	AAU11781	Spider natural sil	193	7	1.4	69	21	AAV61636	Arabidopsis thalia
121	8	1.6	661	22	ABB63580	Drosophila melanog	194	7	1.4	69	21	AAV39788	Arabidopsis thalia
122	8	1.6	674	22	AAB94461	Human protein sequ	195	7	1.4	69	21	AAV30494	Human polypeptide
123	8	1.6	674	22	AAB95164	Human protein sequ	196	7	1.4	69	22	AAO10885	Mutant peptide der
124	8	1.6	674	23	ABG66691	Human novel polype	197	7	1.4	70	21	AAV18678	N55 region from do
125	8	1.6	691	22	ABB58013	Drosophila melanog	198	7	1.4	75	14	AAR37924	N55 region from do
126	8	1.6	692	22	ABG22381	Novel human diagno	199	7	1.4	75	14	AAR37925	N55 region from do
127	8	1.6	718	12	AAR14308	N. clavipes draglin	200	7	1.4	75	22	AAV53732	Human colon cancer
128	8	1.6	718	19	AAV53346	Nephila clavipes s	201	7	1.4	76	17	AAR91273	Elmiera gametocyte
129	8	1.6	718	21	AAV59070	N. clavipes spider	202	7	1.4	76	20	AAV27190	Amino acid fragmen
130	8	1.6	718	22	ABB66734	Drosophila melanog	203	7	1.4	76	22	AAU35810	Helicobacter pylor
131	8	1.6	729	22	ABB66224	Drosophila melanog	204	7	1.4	82	21	AAV33679	Arabidopsis thalia
132	8	1.6	765	22	ABB65864	Drosophila melanog	205	7	1.4	85	14	AAR37923	N55 region from do
133	8	1.6	773	22	ABG15668	Novel human diagno	206	7	1.4	87	22	AAU20918	Human novel foetal
134	8	1.6	783	23	AAV49567	Actinoplanes sp ac	207	7	1.4	88	22	AAU32415	Novel human secret
135	8	1.6	868	22	ABB63905	Drosophila melanog	208	7	1.4	89	21	AAB25862	48kD i-antigen re
136	8	1.6	874	22	AAB68459	Drosophila melanog	209	7	1.4	89	21	AAB25863	Arabidopsis thalia
137	8	1.6	925	23	AAO14246	Human presenilin e	210	7	1.4	89	21	AAV20189	Arabidopsis thalia
138	8	1.6	984	22	AAB79887	Corynebacterium gl	211	7	1.4	92	22	AAV90144	C. glutamicum prote
139	8	1.6	984	22	AAB79930	Corynebacterium gl	212	7	1.4	94	22	AAV80047	Corynebacterium gl
140	8	1.6	1043	18	AAV18306	Photobacterium lum	213	7	1.4	96	22	ABG10434	Novel human diagno
141	8	1.6	1043	19	AAV56569	Toxin TccC, encode	214	7	1.4	97	21	AAV30493	Arabidopsis thalia
142	8	1.6	1083	22	AAV93160	C. glutamicum prote	215	7	1.4	97	21	AAV55019	Arabidopsis thalia
143	8	1.6	1138	21	AAV83222	CaP6 polypeptide (	216	7	1.4	97	21	AAV57951	Human transmembran
144	8	1.6	1195	22	ABB71157	Drosophila melanog	217	7	1.4	98	21	AAV33678	Arabidopsis thalia
145	8	1.6	1588	23	ABB09437	H. influenzae DXR	218	7	1.4	100	20	AAV48556	Human breast tumou
146	8	1.6	1679	22	AAU07343	1-aminocyclopropan	219	7	1.4	100	21	AAV16395	Arabidopsis thalia
147	8	1.6	1698	22	ABB65922	Drosophila melanog	220	7	1.4	100	21	AAV19787	Arabidopsis thalia
148	8	1.6	1700	21	AAB18144	Plasmodium falcipla	221	7	1.4	101	16	AAR98054	Spider dragline va
149	8	1.6	2441	22	ABB62231	Drosophila melanog	222	7	1.4	101	16	AAR98056	Spider dragline va
150	8	1.6	5532	21	AAB23752	S. avermitilis ave	223	7	1.4	101	16	AAR99052	Spider dragline va
151	8	1.6	5532	22	AAG65267	Streptomyces averm	224	7	1.4	101	23	AAU11790	Dragline protein 1
152	7	1.4	7	21	AAB08325	Epitope derived fr	225	7	1.4	101	23	AAU11791	Dragline protein 1
153	7	1.4	9	21	AAB08332	Epitope derived fr	226	7	1.4	102	16	AAR80088	Human derived heav
154	7	1.4	17	22	AAB97243	Gol9 linker librar	227	7	1.4	102	20	AAV95486	Human derived rR3
155	7	1.4	21	19	AAV53348	Nephila clavipes s	228	7	1.4	102	21	AAV44758	Spider silk reseemb
156	7	1.4	21	21	AAV59072	N. clavipes spider	229	7	1.4	103	18	AAV28093	Amino acid sequenc

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:32:30 ; Search time 31 Seconds  
(without alignments)  
2132.010 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496  
Sequence: 1 DTNMSISSSGPDKNKNMS.....LKAYAAISGAIGAKHTNMF 496

Scoring table: OLIGO  
Gapop 60.0 , Gapect 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

A.Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	100.0	496	21	AA083322
2	487	98.2	496	20	AAV35449
3	386	77.8	649	17	AA094584
4	285	57.5	488	17	AA094579
5	259	52.2	259	17	AA094586
6	259	52.2	259	18	AA01743
7	259	52.2	271	17	AA094580
8	259	52.2	302	21	AA083323
9	259	52.2	432	17	AA094585
10	14	2.8	14	21	AA083335

11	14	2.8	14	21	AA083336	Epitope derived fr
12	11	2.2	1937	22	AB058985	Drosophila melanog
13	10	2.0	129	23	AB031472	Human ORF45, prote
14	10	2.0	720	23	AA014994	Laminin-related pr
15	10	2.0	755	22	AA050055	Protonibacterium
16	9	1.8	9	21	AA083326	Epitope derived fr
17	9	1.8	9	21	AA083326	Epitope derived fr
18	9	1.8	9	21	AA083330	Epitope derived fr
19	9	1.8	9	21	AA083333	Epitope derived fr
20	9	1.8	9	21	AA083334	Epitope derived fr
21	9	1.8	111	16	AA072659	Cladosporium herba
22	9	1.8	118	21	AA011695	Arabidopsis thalia
23	9	1.8	149	21	AA011694	Arabidopsis thalia
24	9	1.8	153	22	AB038670	Peptide #6176 enco
25	9	1.8	153	22	AB023746	Protein #5745 enco
26	9	1.8	153	22	AA059299	Human brain expres
27	9	1.8	153	22	AA071846	Human bone marrow
28	9	1.8	153	22	AA019335	Peptide #5769 enco
29	9	1.8	153	22	AA032131	Peptide #6168 enco
30	9	1.8	153	23	AA041659	Human peptide enco
31	9	1.8	159	21	AA011693	Arabidopsis thalia
32	9	1.8	169	20	AA042438	Amno acid sequenc
33	9	1.8	169	21	AA011696	Human peripheral-t
34	9	1.8	169	21	AA093604	Protein encoded by
35	9	1.8	172	21	AA038041	Fragment of human
36	9	1.8	354	22	AB005860	Novel human diagno
37	9	1.8	379	21	AA010364	Arabidopsis thalia
38	9	1.8	382	21	AA010363	Arabidopsis thalia
39	9	1.8	477	23	AB092256	Herbicideally activ
40	9	1.8	481	20	AA037632	Protein which is s
41	9	1.8	513	22	AA078562	Human protein SEQ
42	9	1.8	518	22	AA079546	Human protein SEQ
43	9	1.8	555	21	AA012164	Hydrophobic domain
44	9	1.8	555	22	AB073670	Human oxidoreducta
45	9	1.8	696	22	AB058615	Drosophila melanog
46	9	1.8	1001	22	AB058298	Drosophila melanog
47	9	1.8	1024	22	AB061798	Drosophila melanog
48	9	1.8	1615	22	AA081211	Mycobacterium tbe
49	8	1.6	8	21	AA083328	Epitope derived fr
50	8	1.6	27	21	AA051199	Human SCA7 protein
51	8	1.6	28	21	AA051913	N. clavipes spider
52	8	1.6	30	21	AA059124	N. clavipes spider
53	8	1.6	33	19	AA053349	Nephila clavipes s
54	8	1.6	33	21	AA059073	N. clavipes spider
55	8	1.6	39	22	AB030746	Peptide #3397 enco
56	8	1.6	39	22	AB035923	Peptide #3429 enco
57	8	1.6	39	22	AB021332	Protein #3331 enco
58	8	1.6	39	22	AA069104	Human brain expres
59	8	1.6	39	22	AA069104	Human bone marrow
60	8	1.6	39	22	AA029425	Peptide #3374 enco
61	8	1.6	39	22	AA004633	Peptide #3315 enco
62	8	1.6	39	22	AA004633	Human peptide enco
63	8	1.6	39	22	AB038701	Novel human diagno
64	8	1.6	81	22	AB025327	Lactobacillus amyl
65	8	1.6	84	18	AA019263	Human polypeptide,
66	8	1.6	109	22	AA093342	C. herbarum allerg
67	8	1.6	111	17	AA006538	Zea mays protein f
68	8	1.6	122	21	AA043947	Zea mays protein f
69	8	1.6	122	21	AA043950	Zea mays protein f
70	8	1.6	129	21	AA051200	Human SCA7 protein
71	8	1.6	138	21	AA043946	Zea mays protein f
72	8	1.6	138	21	AA043949	Zea mays protein f
73	8	1.6	164	20	AA044018	Human glutaredoxin
74	8	1.6	164	21	AA012767	Human glutaredoxin
75	8	1.6	165	21	AA043945	Zea mays protein f
76	8	1.6	165	21	AA043948	Zea mays protein f
77	8	1.6	179	21	AA024333	Arabidopsis thalia
78	8	1.6	179	21	AA046905	Arabidopsis thalia
79	8	1.6	197	23	AA017565	M. catarrhalis WCAL
80	8	1.6	209	22	AB011564	Human secreted pro
81	8	1.6	211	22	AA040499	Arabidopsis thalia
82	8	1.6	213	22	AB059343	Drosophila melanog
83	8	1.6	221	22	AA092753	C. glutamicum prote

```

Db 1058 AAAAAGCTGACGAGCGCGCAAAATTAATTCATTCGCTGACCGCGCTGACCGCGCTAT 1117
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 1118 GCACAGCTGAAGCC-----GGG 1135
QY 123 ILeGluAlaAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerAla 142
Db 1136 GTAGAACAGCGCGGAAAGACGACAGCGGCAAGAGCCCTTAATTAAGCCACGCGAT 1195
QY 143 AlaGlnMetLys-----GluValGluAlaValAlaAlaAlaLeuSer 157
Db 1196 GCGACGCTTAAGACGACGACAGACGCAAGCAAGCAAGCCGAGAAACCGATACATTCTG 1255
QY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177
Db 1256 ACCAAATTCAGGGAACGCTATGCGCTTCAGAAATCAGCTTCCAGGGGTGACAG 1315
QY 178 ProArg---SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
Db 1316 GATATCTGTCAATATCGCGCGCTCCTACTATGCTCATGGCATGTTATTGATGATTG 1375
QY 197 GlyGluAlaThrLysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaAlaAspGln 216
Db 1376 GGCAGAAATATACGAGAAAGACCTGCAAAACGATCTTGCGCTT----- 1417
QY 217 ThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluTyr 236
Db 1418 TTCACGCGCTTGACGAGAGGCGCTCAGCGC---GAGATGGAGAAAGAAATCGCGTAATTC 1474
QY 237 GlnGluMetLysAlaIleGluGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
Db 1475 CAGGAA-----GAGACCGCGAAGCGAGAGAACACACCCCATTTATGGCATGTATC 1525
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 1526 GGGAAAGTCTCGCGCGCGT-----CTAACCATTTGTCAGCGTTGCGCGCTGTT 1576
QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAla 296
Db 1577 TTTACCGGTGGCGGAGTGTGGCGGTGGCGGTGGAGTTCGGTA----- 1624
QY 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316
Db 1625 -----ATGCTGCCGATGAATT----- 1642
QY 317 AlaValAlaGlnAlaValLysGlnAlaVal---IleThrAlaValArgGlnAlaIleThr 335
Db 1643 -----GTGAAGGCGGAGCGAGGTGCTTTATTCAGACGCGCTAAAC 1687
QY 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
Db 1688 CCGATTATGGAGCATGTG-----CTGAAGCCGTTAATGGAG 1723
QY 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
Db 1724 CTGATTGGCAAGCGATTAACCAAGCGCTG-----GAGGATTAGCGCTC 1768
QY 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTyrVal 395
Db 1769 GATAAGAAACGACAGATGGCGGACGATTTGTCGATTTGCGCGCATTTGCC 1828
QY 396 ThrValGlyValGlyValVal-----AlaAlaProAlaLeuGly 409
Db 1829 ATGTGTGGCGCTCATTTGTGTGCGCATTTGTCGGAAGCGCGGCAAACTGGGT 1888
QY 410 LysGlyIleMetGlnMetGlnLeuSerGlnMetGlnLysValAlaGlnPheGlnLys 429
Db 1889 AACGCGCTGAGCAAAATG-----ATGGCGAAACGATTAAAGAGTTGGTGCT 1936
QY 430 GluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTyrGln 449

```

```

Db 1937 AACGTGCTGAAA---CAGTTGGCGGCAAAACGACGCAAACTCTTTACCCAGGGGATGCA 1993
QY 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyIle 461
Db 1994 CGATTACTAGCGGCTGTGGTAAATGTGGGTAGCAAGATGGCCTGCAGAACGATGCTTA 2053
QY 462 SerAsnGlnMet-----ThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLys 478
Db 2054 AGTAAGAGCTGTGTAGTAAATACCTAAATTAAGTGGCGTTGGCATGGAAGTCAAGCAAT 2113
QY 479 AlaTyrAlaAlaIleSerGlyAlaIleAla 488
Db 2114 ACCGACCCAGTCAGCGCGGTGTTGCC 2143

```

Search completed: January 27, 2003, 15:06:56  
Job time : 3020 secs

[illegible]





/note="Similar to proteins from bacteria, archaea and eukaryotes e.g. Entamoeba histolytica acetyl-CoA synthetase TR:09NAT4 (EMBL:AF286346) (713 aa) fasta scores: E(): 0, 34.1% id in 721 aa and Glardia lamblia acetyl-CoA synthetase TR:091IN2 (EMBL:AF107206) (726 aa) fasta scores: E(): 0, 30.0% id in 700 aa orthologue of E. coli yf10-ECOLI; Fasta hit to yf10-ECOLI (886 aa), 92% identity in 886 aa overlap"

/codon\_start=1  
/transl\_table=11  
/product="putative acyl-CoA synthetase"  
/protein\_id="CAD05835.1"  
/db\_xref="GI:16503810"  
/db\_xref="SPTREMBL:082414"  
/translation="MSOGLFALLRPKSIAYVAGSMKPPHRYGIMRNLLAGGNGPVLPVTPAKKAVGVAMPDIALSPPTPDIAIICTNARNLALLDALGAKCKTCTILSAPTSOHEELASARHYKMRLLGPSNIGLPAWGLNAPSVPVTKGCKLAISQSAVSNTILDMOOREMGFSYFLAGDSIDIDVELLDYLRDSTSAIILYLEQLSDARFVSAARSARKPLIVIKSGRSPAAQRLLNTSAGMDPMDAIOAGILLRVDYHLEFSAVETLSHMRPLHGDRLMITSNGAPALADELMSRNGKLATSEETCLOTRTLEPAHIDIANPLDLCDDASSSEHYVKTIDILASODFDLAMYIHSPTAAGTESHAHLETKRHPKGFVLTLLTWGCEPFSQSEARRLFESEAGLPTTTPBSTTAFPMHVEYRNQOLKEPPLPSNLTSTWTAHNLQRAIEGATSLDHEVQPLIAYGLHTLPTWIASOAEAVHAEIOIGYPALKLRSPDIPEKSEVQVMLYETASVEQQAANAIEDRVKAMPQARIGHGLVOSMANRAGAEELRVVEHDPVPGPLIMGEVGEVPEEOAVVALPLMNLARYLVIOGIKORIRARAGLRPLDIVGLSGLVSNLTVDCPELORDILHLASA SERTALDYTDIAPPDGNSRLRVPVPHOLEEYVNGRCRPLPLPDEPOLR OPTAOTKEDLYRRYRSEINETHEDIANMTQIDVDREAFVAYRMDNAEETILEYTRALSDPDNDAEFAVLRSDKGLGJRRLMEKLIAYTRDHGLKRLNGITPNNKGVAA LARKLGFQVDYIDLDGIVGLTINLAKCES"  
/gene="stry2844"  
/note="Pfam match to entry PF00583 Acetyltransf. acetyltransferase (GNAT) family, score 59.00, E-value 9.9e-14  
6255..7610  
/gene="stry2845"  
/note="psa"  
6255..7610  
/gene="stry2845"  
/BC\_number="2.7.8.8"  
/note="Orthologue of E. coli psaA (PSS\_ECOLI); Fasta hit to PSS\_ECOLI (451 aa), 94% identity in 451 aa overlap"  
/codon\_start=1  
/transl\_table=11  
/product="CDP-diacylglycerol-serine O-phosphatidyltransferase"  
/protein\_id="CAD05836.1"  
/db\_xref="GI:16503811"  
/db\_xref="SPTREMBL:08XFM5"  
/translation="MLSKFRNKHOOHLOALPKISQSVDDVDFYTPATPREFLEKI ASAPORICVAYILEDDOGKGLDALYAAKQREBELRVYVDHMRQRGITGAASNTNADWCTGLAENPGLIDVPYGVPLNREALGVLFKGFIIIDSVLVSASGLNLYLHGHDRYRDRYOLIRNRQADIMFPMVONLNGKVNLDLDFORPSEIENDIRLYRQELSDASYHFGDANDEQLSVTPLVGLKSSLLKTFHLMPCAEHKLITCTPYENLPAVLRNITFOLLRDKKVEIIVGDGTANDEFYIPEEPFETIGALPIYVETNRPFSRLQYVNTQOLVYRLKMDNDNTYHLKGMWDDKMLITGNNLIRPMARLDLEAIIID PKQELAPQREKLELIRHTTIVKHNRDGLSTADIPIKYKRLIRLRIRIDRLISRI L"  
/gene="stry2845"  
6651..6731

misc\_feature  
6651..6731  
/gene="stry2845"  
Alignment Scores:  
Pred. No.: 0.447  
Score: 220.50  
Percent Similarity: 41.02%  
Best Local Similarity: 23.67%  
Query Match: 9.31%  
DB: 1  
Gaps: 20

US-09-889-314-2 (1-496) x AL627276 (1-274050)

28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnLthrLysGlnIleGlnGlnIthr 47

Db 180562 CAGCGCATGATGAGTCACACAAAAGAGATGGGATTCAGGTATCGAAAGAAATTCACGACG 180503  
||| : : : ||| : : : |||||  
Qy 48 ArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGlyLys 67  
||| : : : ||| : : : |||||  
Db 180502 GCTCTGGGA-----GAGGCTCAGGAGGCGACGAGATCTATGAAACCGACGATCAAA 180452  
||| : : : ||| : : : |||||  
Qy 68 AspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnIleAlaGly 87  
||||| : : : |||||  
Db 180451 -----AAGACGATACCGCCCAAGCTTTATACGCTGGGCC 180413  
||| : : : ||| : : : |||||  
Qy 88 LysGluSerSerGluSerGlnLysAla-----GlyAlaAspThrGlyVal 102  
||||| : : : ||| : : : |||||  
Db 180412 AAAAAGCTGACGAGCGGCAAAATTAATTGCAATGCTGACCCACTGATACCCGGCTAT 180353  
||| : : : ||| : : : |||||  
Qy 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122  
||| : : : ||| : : : |||||  
Db 180352 GCACAAGCTGAACCC-----GGC 180335  
||| : : : ||| : : : |||||  
Qy 123 IleGlnGlnAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerAla 142  
||| : : : ||| : : : |||||  
Db 180334 GTAGAACAGCGCGAAGAAAGACGACAGCGCAAGAGCGCTTAGATTAAGCCACGCAAT 180275  
||| : : : ||| : : : |||||  
Qy 143 AlaGlnMetLys-----GluValGluAlaValValAlaAlaLeuSer 157  
||| : : : ||| : : : |||||  
Db 180274 GCGACGGTTAAAGCAGCAGCACAGACGCAAGAGCGCAAGAGCGGATTAACATTCTG 180215  
||| : : : ||| : : : |||||  
Qy 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThr 177  
||| : : : ||| : : : |||||  
Db 180214 ACCAAATTCAGGCAAGCGCTATGCGCGCTCAGAAATCAGGTTCCAGAGGTGAGCAG 180155  
||| : : : ||| : : : |||||  
Qy 178 ProArg-----SerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196  
||| : : : ||| : : : |||||  
Db 180154 GATATCTGTCATATGTCGCCGCCCTCACTATGCTATGCGCATGTTTATTGATGATTG 180095  
||| : : : ||| : : : |||||  
Qy 197 GlyGlnAlaThrLysSerAlaLeuSerAsnTyraLaserThrGlnAlaGlnAlaAspGln 216  
||| : : : ||| : : : |||||  
Db 180094 GCGAAATATACGAGCAAGAAAGCGTCAAAAGCATCTGCGCTT----- 180053  
||| : : : ||| : : : |||||  
Qy 217 ThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlyThr 236  
||| : : : ||| : : : |||||  
Db 180052 TTCACGCGCTTGACGAGAGGCGCTCAGGCG--GAATGGAAGAAATCGCGTGAATTC 179996  
||| : : : ||| : : : |||||  
Qy 237 GlnGlnMetLysAlaIleGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal 256  
||| : : : ||| : : : |||||  
Db 179995 CAGGAA-----GAGACGCCCAAGCCGAGAAACGACCGCATATGCGATGATC 179945  
||| : : : ||| : : : |||||  
Qy 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276  
||| : : : ||| : : : |||||  
Db 179944 GCGAAAGTCTCTCGCGCGCTG-----CTAACCATTTGTACGCTTGCGCGCTGTT 179894  
||| : : : ||| : : : |||||  
Qy 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAla 296  
||| : : : ||| : : : |||||  
Db 179893 TTACCGGCGGCGGCGTCTGCGCGCTGCGCGCTGCGGCTTGGCGTTA----- 179846  
||| : : : ||| : : : |||||  
Qy 297 AlaGlyAlaAlaIleGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316  
||| : : : ||| : : : |||||  
Db 179845 -----ATGTGGCCATGAATT----- 179828  
||| : : : ||| : : : |||||  
Qy 317 AlaValValGlnAlaValLysGlnAlaVal-----IleThrAlaValArgGlnAlaIleThr 335  
||| : : : ||| : : : |||||  
Db 179827 -----GTGAAGCGCGGCGGCGGTGCTGTTATTCACGACGCGCTAAGC 179783  
||| : : : ||| : : : |||||  
Qy 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355  
||| : : : ||| : : : |||||  
Db 179782 CCGATTATGAGCATGTG-----CTGAAGCGCGTTAAAGGAG 179747  
||| : : : ||| : : : |||||  
Qy 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375  
||| : : : ||| : : : |||||  
Db 179746 CTGATTGGCAAGCGCATTTACCAAGCGCTG-----GAAGCATTTAGCGCTC 179702  
||| : : : ||| : : : |||||  
Qy 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrVal 395  
||| : : : ||| : : : |||||

[illegible]



QY 410 LysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLys 429  
Db 1857 AACGGCGTGAAGAAATG-----ATGGCGCAACCATTAAGACGTGGTGCT 1904  
QY 430 GluValGlyLysLeuGlnAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln 449  
Db 1905 AACGGCTGAAA---CAGTTGGCACAAACGCGACCAACTCTTACCAGGGGATGCA 1961  
QY 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyGlu 461  
Db 1962 CGTATTACTAGCGCGCTGGTAATGTGGTACAGATGGCGCTCAACGAATGCTTA 2021  
QY 462 SerAsnGlnMet-----ThrGlnLysAlaThrLysGlyAlaGlnIleLeuLys 478  
Db 2022 ACTAAGAAGCTGTGTAATATCCCTAATAAATAGTGGCTTGCGATGGAATGCAGAT 2081  
QY 479 AlaTrpAlaAlaIleSerGlyAlaIleAla 488  
Db 2082 ACCGACGCCACTCAGCGGTGGGTGCC 2111

RESULT 41  
AR082442 AR082442 5393 bp DNA linear PAT 31-AUG-2000  
LOCUS AR082442 Sequence 9 from patent us 5972899.  
DEFINITION AR082442  
ACCESSION AR082442 GI:10009168  
VERSION AR082442.1 GI:10009168  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5393)  
AUTHORS Zychlinsky, A. and Chen, Y.  
TITLE Apoptosis induced by Shigella IPaB  
JOURNAL Patent: US 5972899-A 9/26-Oct-1999;  
FEATURES  
source 1. 5393  
Location/Qualifiers  
/organism="unknown"  
BASE COUNT 1584 a 1156 c 1458 g 1195 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 0.00397 Length: 5393  
Score: 220.50 Matches: 116  
Percent Similarity: 41.02% Conservative: 85  
Best Local Similarity: 23.67% Mismatches: 182  
Query Match: 9.31% Indels: 107  
DB: Gaps: 20

US-09-889-314-2 (1-496) x AR082442 (1-5393)

QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThr 47  
Db 876 CAGGCGATGATGATGATCAAAAGAGATGGGATTCAGATTCGAAGAATTCACAGCG 935  
QY 48 ArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyLysSerGlyLys 67  
Db 936 GCTGTGGG-----GAGGCTCAGAGGCGACGATCTCTATGAAACCCAGCATCAAA 986  
QY 68 AspLysThrSerThrThrLysThrGlnThrAlaProGlnGlnGlyAlaIleAlaGly 87  
Db 987 -----AAGACGATACCGCCCAAGAGCTTTATGACCTGCGGCC 1025  
QY 88 LysGlnSerSerGlnSerGlnLysAla-----GlyAlaAspThrGlyVal 102  
Db 1026 AAAAAAATCAGCAGCGCAAAATTAATTCATGCTGAGCCAGCTGACCCCGGCTAT 1085  
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122  
Db 1086 GCACAGCTGAAGC-----GGC 1103  
QY 123 IleGlnGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuSerLeuSerAla 142  
Db -----GGC 1103

Db 1104 GTAGAACAGCGCGGAAAAAGACAGAGCGAAAGAGCGCTTATGATTAAGCCAGCAT 1163  
QY 143 AlaGlnMetLys-----GluValGluAlaValAlaValAlaLeuSer 157  
Db 1164 GCGACGGTTAAAGCGAGCAGACAGCCCAAGAGCGAAAGCCGAAAGCGGATTAATTCG 1223  
QY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177  
Db 1224 ACCAATTCACAGGAAGGCTATGCGCCCTCTCAGATTCAGATTCAGGTTCCAGGGTGACG 1283  
QY 178 ProArg---SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196  
Db 1284 GATTAATCTCAGAAATGTGCGCCCGCTCATATGCTCAGAGGCGATGTTATGAGTTGTG 1343  
QY 197 GlyGluAlaThrLysSerAlaLeuSerAsnTrpAlaSerThrGlnAlaGlnAlaAspGln 216  
Db 1344 GCGAAAAATACGGAAGAAAGCTCGCAAAACGATCTTGGCGTT----- 1385  
QY 217 ThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGluTrp 236  
Db 1386 TTCACGCCCTTCACAGAAAGGCGCTCAGCG---GAGATGGAAGAAATTCGCGCTGAATTC 1442  
QY 237 GlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal 256  
Db 1443 CAGGAA-----GAGACGCGCAAGCCGAGAAACGACCATTAATGAGATGATTC 1493  
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276  
Db 1494 GGGAAAGTCTCTGCGCGCTG-----CTAACCATTTGCTAGCGTTGCGCTCGT 1544  
QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGlyAlaAlaAla 296  
Db 1545 TTTACCGGTGGGCGCAGTGTGGCGCTGCGGTGGCGGACTTCGCGGTA----- 1592  
QY 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316  
Db 1593 -----ATNGCGCGCGATGAATT----- 1610  
QY 317 AlaValAlaGlnAlaValLysGlnAlaVal---IleThrAlaValArgGlnAlaIleThr 335  
Db 1611 -----GTGAAGCGGCGAGCGGGGTGCTGTTATTCAGACGCGCTAAAC 1655  
QY 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355  
Db 1656 CCGATTATGACACATGT-----CTAACCGCTTATTCGAG 1691  
QY 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375  
Db 1692 CTGATTGGCAAGCGATTAACCAACCGCTG-----GAGGATTTAGCGCTC 1736  
QY 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpVal 395  
Db 1737 GATTAAGAAAAACGACAGATGCGACGACGATGTTGGTATTCGATGTCGCCCTATTGCC 1796  
QY 396 ThrValGlyValGlyValValVal-----AlaAlaProAlaLeuGly 409  
Db 1797 ATGTAGCGGCTATGTTGTGTGTGTCAGTTGCGGAAAGCGCGCGCAACTGGGT 1856  
QY 410 LysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLys 429  
Db 1857 AACGGCGTGAAGAAATG-----ATGGCGCAACCATTAAGATGGTGCT 1904  
QY 430 GluValGlyLysLeuGlnAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln 449  
Db 1905 AACGGCTGAAA---CAGTTGGCACAAACGCGACCAACTCTTACCAGGGGATGCA 1961  
QY 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyGlu 461  
Db 1962 CGTATTACTAGCGCGCTGGTATGTGGTACAGATGGCGCTCAACGAATGCTTA 2021  
QY 462 SerAsnGlnMet-----ThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLys 478  
Db 2022 ACTAAGAAGCTGTGTAATATCCCTAATAAATAGTGGCGTTGGCATGGAATGCAGAT 2081

```

KAEKADNLTFRQGTANASQVSGEDNLSNVARLTMLAMFLEIYKNTESLQ
NDLAFNLQEGROAEEMKSAEFOETKAEETNRIMCICIKVIGALLTIVSVAA
FTGASIALAAAGVAMVADIEVKATGVSFTIOALNPIMEHIVPRIMELIKATKA
LEGIGYDKRTAMAGSIYGAITAAIAMAIVYVAVGGAALGNALSKMGETIK
KLVPVYKOLAQNSKLTQGMORITSGLVNCSKGLDTNALSELVNTLNKVALG
MEVNTAASAGVAGVFIKNASEALDFMLARFAMDIOOVLKOSVEIFGENOKVT
ALQKAMSSAVOONADASFILROSHA"
2352..3581
/gene="sipc"
/gene="sipc"
2352..3581
/gene="sipc"
/codon_start=1
/transl_table=11
/protein_id="CAA57989.1"
/db_xref="GI:1009004"
/db_xref="SPTREMBL:Q56135"
/translacion="MLISVNGINPAAYLVNHNHSEVSSQTSQSVASKDILNSIGISS
KYVDGLSPTLSAPAPGYLTQPGTITTSFLKASIONTDMNODLMLANNVTKRANVY
OTLREOQAEVGFEDISGSSSAVALLAANTMLTINQADSKLSGKSLVSPDAK
TTASSMMREGMNAISGISOSALOLGITGVAKLEKGLONERGAALKHNAKIDKLT
BSHSIKNVIQNSGVKICAEVYDSLKSLMKRTGDATKTNIDATLKSAGTSASESL
GTRNSKQISPEHQAILSKRLSVESDINLEONTDMTRIDARKMOTGDLIMKNSVT
VGGIAGASROYATQERSEQIISOVNNRVASTADSEARSSRSKSTSLIQEMUKTMEI
NOSKASALAAIIGNIRA"
3651..4670
/gene="snpd"
/gene="snpd"
3651..4670
/gene="snpd"
/codon_start=1
/transl_table=11
/protein_id="CAA57990.1"
/db_xref="GI:1009005"
/db_xref="SPTREMBL:Q56136"
/translacion="MNTIONISASPHPGIVARPPPSASEHAEIAVNDSTEHRCGT
ITLSQAARIDQAOOTLOSTPPISEENDEERTLROQLTSSLMALAKSVLSAEON
ENLRSTFSARRRYALRLMPARTTISDAEIMVDSQNSAISGDSYLGVEVNAVYT
DEYQAFSDLSKMGWMLSPKDGNTIKLVNDSKSISSLINKYQINKNTILFPST
GSGMTATKAEAEOMIKELNPDCLKASGCVYLVDPGIPSKVSDINGISGSAL
EUDNAKYQAWOSGFKAEENLKTITLOTLQTKSNANSLYDNLKYLSTISSLELAK
SFLQG"
4689..5393
/gene="sipa"
/gene="sipa"
4689..5393
/gene="sipa"
/codon_start=1
/transl_table=11
/protein_id="CAA57991.1"
/db_xref="GI:1009006"
/db_xref="SPTREMBL:Q56137"
/translacion="AVTSVROPPYIAMPQNOTEIKTQATNLAAANLSAVRESATLTS
ETKGOLEDFPALIKOASIDALFKCGKDEALKVEVTNNVACGKKAIMEFAGIPPSA
LNATSDSPARTLLMKVCAEYTAQIITKDLKKSAGPWLPTKKAELALENEKOL
DITKNTGESLSTNLVMOEVMPIVASCIEHNFCTLDPLRSTLQVLVKAAGA
VEALDMCHOKLQEOGT"
BASE COUNT 1584 a 1156 c 1458 g 1195 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00397 Length: 5393
Score: 220.50 Matches: 116
Percent Similarity: 41.02% Conservative: 85
Best Local Similarity: 23.67% Mismatches: 182
Query Match: 9.31% Indels: 107
DB: 1 Gaps: 20
US-09-889-314-2 (1-496) x STSIPGNS (1-5393)
QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThr 47
DB 876 CAGGCGATGATTGATGCACAAAAAGAGATGGGATTCAGGTATCGAAAGATTCCAGAGC 95
QY 48 ArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLys 67

```

```

DB 936 GCTGTGGGA-----GAGGCTCAGAGAGCGACGATCTATGAGCCAGCATCAAA 986
QY 68 AspLysThrSerSerThrThrLysThrGluThrAlaProGlnIleGlyValAlaIleGly 87
DB 987 -----AGACGGATVACCGCAAGAGTGTATGACGCTGGCGCC 1025
QY 88 LysGluSerSerGluSerGlnLysAla-----GlyAlaAspThrGlyVal 102
DB 1026 AAAAACTGACGACGCGCAAAATAATTCATCGCTGAGACCCAGCTGACCCCGCTAT 1085
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
DB 1086 GCACACCTGAAGCC-----GCC 1103
QY 123 IleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAla 142
DB 1104 GTGAAAGAGCGCGCAAAAGAACGACAGAGCGCAAGAGCCCTTAGTAAGCCAGGAT 1163
QY 143 AlAGlnMetLys-----GluValGluAlaValValAlaAlaAlaLeuSer 157
DB 1164 GCGACGCTTAAGCAGGACGACAGACCCCAAGACCAAGCCGAGAAAGCGGATTAACATTCTG 1223
QY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177
DB 1224 ACCAAATTCAGGGAACGGCTATGCGCGCTCAGATCAGGTTTCCAGCTTGAGCAG 1283
QY 178 ProArg-----SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
DB 1284 GATATCTGTCAAATGCTGCGCCGCTCACATATGCTCATGCTTATGAGATGTG 1343
QY 197 GlyGluAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGln 216
DB 1344 GCGAAATAATCGGAGAGAAACCTGCAAAACGATTCGCTT----- 1385
QY 217 ThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGluTyr 236
DB 1386 TTCACGCGCTTGACAGGAAGGCGTCAAGCGC--GAGATGGAAGAAATCGCTGAATTTC 1442
QY 237 GlnGluMetLysAlaAlaGlnLysLysSerLysAspLeuGlnGlyThrMetAspThrVal 256
DB 1443 CAGCAA-----GAGACGCGCAAAAGCCGAGAAACGACGATTAAGGATGATC 1493
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
DB 1494 GCGAAAGTCTCGCGCGCGCTG-----CTAACCATTTCTCAGCTGTGGCGCTGTT 1544
QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAla 296
DB 1545 TTTCACGCTGGGCGAGTGTGCGCGTGGCTGCGGTGGACTTCGGA----- 1592
QY 297 AlaGlyCysAlaAlaGlyAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGln 316
DB 1593 -----ATGGTGCCATGAAATTT----- 1610
QY 317 AlaValValGlnAlaValLysGlnAlaVal-----IleThrAlaValArgGlnAlaIleThr 335
DB 1611 -----GTGAAGCGCGGACGCGGGGTGCTTTATTACGACGCGCTTAAC 1655
QY 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
DB 1656 CCGATTATGAGCATGNG-----CTGAAGCCGTTAAATGAGG 1691
QY 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
DB 1692 CTGATTGGCAAGCGATTAACAAAGCGCTG-----GAGGATTAGGCGTGC 1736
QY 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTyrVal 395
DB 1737 GATTAAGAAAAGCGCAGAGATGCGACGACGATTTGTGTGCTGTGCGCTATTGCC 1796
QY 396 ThrValGlyValGlyValValVal-----AlaAlaProAlaLeuGly 409
DB 1797 ATGCTAGCGGCTCATTTGGTGGTCTGCGAGTTGTGCGGAAGAGCGCGCGCGCAATCGGGT 1856

```

```

Db 174863 AATAGCAGCAGTAACAAACGACCAACAGCAATTAAGCAGCAGCAGCAACAC 174804
Oy 101 yValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThr-----AlaThrIysII 117
Db 174803 AACAGTAGCAGCAGCAGCAACCAACCAATTAACAAAAGCAACGATAGCAACATATC 174744
Oy 117 eAlaMetGlnThrSerIleGluValaSerIleSerMetGluSerThrIleuGluSerIe 137
Db 174743 AGCAACAGCAACGCAATTAAGCAACAGCAGCAGTCA-----GCACATATACAGCAGT 174690
Oy 137 uGlnSerIleuSerAlaAlaGlnMetIleuSerGluValaGluAlaValaValaAlaAlaLeuSe 157
Db 174689 AAAAACAGCAACGACGACGACCAACACAGCAACAGCAACATTAAGACCA---GC 174633
Oy 157 rGlyIysSerSerGlySerAlaIleuGluIleuThrProGluLeuProIyProGlyValIh 177
Db 174632 AGCAACAGCAACGATAGCAGCAGCAGCAACACAGCAACATTAACAAAGCAACGATAGC 174573
Oy 177 rProAtgSerGluValIleGluIleGluIleuAlaAlaAlaIleGlnThrIleuG 197
Db 174572 AACGATATCA-----GCACACGACGACGCAATTAAGCAACAAAGC 174534
Oy 197 yGluAlaThrIleuSerAlaLeuSerAsnTyraIaSerThrGlnAlaGlnAlaAspIlnTh 217
Db 174533 ATCAGCAACAGCAATTAAGCAGCAACATTAACCAACAGCAACATTAAGCAGCAACAGCAAC 174474
Oy 217 rAsnIleuGlyLeuGluIleuGluIleuAlaIleuIleuIleuIleuIleuIleuIleuIleu 237
Db 174473 AGCAGCAGCAGCAACGAAATTCAGTAAATTAACAGTA----- 174437
Oy 237 nGluMetIleuAlaAlaGluIleuIleuSerIleuAspLeuGluGluIleuMetAspThrValas 257
Db 174436 -----GCACGACGACGACGACATTCATACCATTAAGTACCAACACAGTACAGCAAC 174387
Oy 257 nThrValMetIleuAlaValaSerValaAlaIleuThrValIleSerIleuValaAlaAlaIleu 277
Db 174386 ACCAGCAGCAATAGCAGCAGCAGCAATAGCA---ACAGTAATAGTACAGCAACAGCAATAGC 174330
Oy 277 eThrIyGlyAlaGlyLeuAlaGlyLeuAlaGlyLeuAlaGlyAlaValaGlyAlaAlaAlaAl 297
Db 174329 AACAGTAGCAGCAGTCAACCAACAGCAACAGCAGCAACCAATTAAGCAACCCAGC 174270
Oy 297 aGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 317
Db 174269 AACAGCAACAGTACGACGACGACCAACACAGTACGACGACGACCAACACAGCAAC 174210
Oy 317 aValValGlnAla-----ValIy 323
Db 174209 AATTAACAAAGCAACAAATAGCAACAGTATACGACCAACAGCAATAGCAGCATCAGCAAC 174150
Oy 323 sGlnAlaValIleuThrAlaValaIaArgIlnAlaIleuThrAlaAlaIleuIleuAlaAlaValIy 343
Db 174149 AACAGCAACAAATAGCAGCAGTAAACAGCAACAGCAACAGCAACCAACAGCAGCAGC 174090
Oy 343 sSerGlyIleuIleuAlaPheIleuThrIleuValIleuValaIleuAlaIleuAlaAlaAlaAla 362
Db 174089 AGCAACAAATTAAGCAGCAACCAACAGCAACAGTACGACGACGACGACCAACATTAAG 174030
Oy 362 rIlyGlyIleuIleuSerIleuValaPheAlaIleuGlyIleuIleuIleuIleuIleuIleuIleu 381
Db 174029 AAAAGCAACAAATAGCAACAGTATACGACCAACAGCAACAGCAACATTAAGCAAAA-----GC 173976
Oy 381 oIyIleuSerIleuValaIleuSerIleuThrIleuIleuIleuIleuIleuIleuIleuIleu 400
Db 173975 ATCAGCAACAGCAATAGCATCAATTAACCAACAGCAACCAATTAAGCAGCAACAGCAAC 173916
Oy 401 -----ValValIleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 411
Db 173915 AGCAGCAGTACGCAACGAATATCACTAATTAACAGTACGACGACGACGACGACGACGACGACG 173856
Oy 411 yIleuMetGlnMetGlnLeuSerGlnIleuSerGlnIleuSerGlnIleuSerGlnIleuSerGlnIleu 431

```

```

Db 173855 AATCATTAACATAGTAGCAACAAACAGTACAGCAACACCAGCAGCAACATAGTAGCAACAT 173796
Oy 431 IgIyIleuGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 451
Db 173795 AGCAACAGTAATAGTAGCAGCAACAGCAACAGCAACAGTACGATCGGCAACAAAGAC 173736
Oy 451 aSerIleuAlaSerIleuGlnIleuGluSerAsnGluMetThrGlnIleuAlaAlaAlaAlaAla 471
Db 173735 AACA---ATAGCAGCAGTAAACACAGCAACAGCA-----GCACCAAC 173697
Oy 471 sIleuGlyAlaGlnIleuIleuIleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 491
Db 173696 AATAGCAGCA-----GCCGACGACAGCAACAGCAACAGTACGACGACGACGACG 173655
Oy 491 sIySThr 493
Db 173654 AACCACA 173648
RESULT 40
STSPGNS
LOCUS STSPGNS 5393 bp DNA linear BCT 04-OCT-1995
DEFINITION S.typhi spat, sipB, sipC, sipD and sipA genes.
ACCESSION X82670
VERSION X82670.1 GI:1009001
KEYWORDS sipA gene; sipB gene; sipC gene; sipD gene; spat gene.
SOURCE Salmonella typhi.
ORGANISM Salmonella typhi.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE 1 (bases 1 to 5393)
AUTHORS Hernant,D., Menard,R., Arlicau,N., Parsot,C. and Popoff,M.Y.
TITLE Functional conservation of the Salmonella and Shigella effectors of
entry into epithelial cells
JOURNAL Mol. Microbiol. 17 (4), 781-789 (1995)
MEDLINE 96111497
PUBMED 8801431
REFERENCE 2 (bases 1 to 5393)
AUTHORS Popoff,M.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1994) M. Popoff, Institut Pasteur, Unite des
Enterobacteries, 28 rue du Docteur Roux, 75724 Paris Cedex 15,
FRANCE
FEATURES
source location/Qualifiers
1..5393
/organism="Salmonella typhi"
/strain="Ty2"
/db_xref="taxon:601"
43..540
/gene="spat"
43..540
/gene="sipB"
/codon_start=1
/transl_table=11
/protein_id="CAA57987.1"
/db_xref="GI:1009002"
/db_xref="SPTREMBL:057260"
/transl_table="MDYNNVSESEVVAEIMDAVESEATIKDVHGIPDDMGILYAH
VEFYNGRIUDEETFFRFLCTIDFYNPDTTGLAAVCOLKQFOACDLYAVATFLK
NDYRPVFTGQCQLMRKAAROCFELVNETTESLRKALVYLEALKTAEYBQHS
EOEKE"
543..2324
/gene="sipB"
543..2324
/gene="sipB"
/codon_start=1
/transl_table=11
/protein_id="CAA57988.1"
/db_xref="GI:1009003"
/db_xref="SPTREMBL:056134"
/transl_table="NYNDASSISRSGLYTONPPLAAREGVAKKNTDFLKAADKAKDV
VATKAGDILAKTGSSESAINTVGLKPPIDAAREKLSSRQTLTLGLKMTLLGDLVSL
QLESRLAVQAMIESOKEMGIQVSEKFOPLAGEADVLDYASIKRTDTAKSVYDAA
AKTLQAKNKLQSLDPADPGYQAQAAVAQAGKEATTEAKFALDKRTADATVKKAGTAKA

```





gene  
CDS  
/gene="TC0868"  
2822..4132  
/gene="TC0868"  
/note="conserved hypothetical protein; identified by  
glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF39664.1"  
/db\_xref="GI:7190895"  
/translation="MTGTGRCGDMTPDPSPLLAQLTQNMANSASADKNGCAAKTQOQEER  
SPEDILIOSOGGKSKASTSOTSKSGSEKAOKSGTTTTSVAOASQPTAAVAG  
AASSFPNDGASLSPSPSNPNVGLKKMGTLALMLITFLAQAASAKSNSSPOO  
ONQALQNVYAMAPETIGNAIRTQANHQAATLQAKOSLISITITVGFAYSVGGITLS  
ASRSIGGKSAFAETNETATPAASASATSAAKTAVNALDVANVAATGAKASGAASAA  
SAAKRLQNMMDIASKTLQSOTASKATGELGNALNTPMSEKISRGLNVAKTQGGRAA  
AOFARLASSAASISOMYHGLAGVADITGGITIGAVNDEORAGMAEARAEELKSLN  
SVQSYAQAQOQLQDSOOSFNSALQTLQSTIADSSLQTTGAMFN"  
complement(4206..5162)  
/gene="TC0869"  
complement(4206..5162)  
/gene="TC0869"  
/note="conserved hypothetical protein; identified by  
glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF39665.1"  
/db\_xref="GI:7190896"  
/translation="MMVLPRANLYVEPMALFLPLNFIWSSSFALSKAMDASDPL  
EYTSGRWLAVGLVLEGLLFRKRSRFLPRQVMPTVLLSVGLITLVNLEPTIGLQGIS  
SKACFLTFGSPFTAFCSYVQLREVYWKRLGSLGVLVLYLLFGGEDVSEWNGS  
WOLGPELLILVATLSSYGYTLRLKLERQCESLTAINVAVINGILSLASATTT  
EWMNLPDENPVLFLQSIGALVIFSNLCYMLFAKLIRFESSTFLSFCNLYMPFASF  
PCWLLGSEFPGLLFAVGFVWLGCRLLYHBEPRQGYLVSE"  
5409..7316  
/gene="TC0870"  
5409..7316  
/gene="TC0871"  
/note="similar to GB:M36593 SP:P18256 PID:143764  
PID:1565235 GB:AL009126; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="threonyl-tRNA synthetase"  
/protein\_id="AAF39666.1"  
/db\_xref="GI:7190897"  
/translation="MHWTCNDEAFLEPGASANDLASKMNSHCFAGLINDEKDE  
STTLQDGVLELWDDPQGRFLFHTSAHLAQVLRMPSAQPTIGPVIDQGFYTD  
EANTLISEEDFPALIEAMAKAIAEKLALISROYFSKRETLAYFSDNPKLELTLEP  
FNLISAYKGEFLLDCRGPHLPSTAPVAFKLLRTSAAYMGDSKSLIRIYIAPFEP  
TKLEKLEHLLQLEEKAKRDHRYLGKLDLFEQOQSACMPFPHRGVYINNALIDYWK  
RLHQHLAGOOLITPOLMSRELMEISGHEWENKNNYTLTVDEEDYAIKPMKPCGMILY  
XYTQIHSRREPRLIATGIHYHRLHSALSGIMAVYRFHDDAKHVLTPROVEEPTIT  
NLYNLPVEIDITHVDALNRTWCGTLOLDMFLPERFLKYNQGEKSTPIMLRAL  
GGAAVYVGLIDHFGDAINRTWCGTLOLDMFLPERFLKYNQGEKSTPIMLRAL  
FQSIIRFLGILLIEHFKGRFPLMSPEHRIITVADRHRAKQELAKFTFOGIGLATYAD  
QASESVKRIANQNMVNYMITIDGRLELEQTOLAVFTRDNRNLIIVSEHFTWLTILE  
ELRSLTSSSL"  
7785..8552  
/gene="TC0871"  
7785..8552  
/gene="TC0871"  
/note="similar to GB:D26185 SP:P37522 PID:467381  
PID:580906 GB:AL009126; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="para family protein"  
/protein\_id="AAF39667.1"  
/db\_xref="GI:7190898"  
/translation="WKTIAVNSFKGTAKTSTTLGLGALAAQYHKARVLLIDPDAQAN  
LTAQGLDPDCDLSAVLVQGEKQISVETIRSIDSSGDLIPADPWLEVEVSGSLAAD

CDS 8492..10399  
 /gene="chrs"  
 /codon\_start=1  
 /transl\_table=11  
 /product="threonyl tRNA Synthetase"  
 /protein\_id="AAC68183.2"  
 /db\_xref="GI:6578108"  
 /translation="MIHYCNQAELEPEGASAMDLANKMKSHCFAGALINDKERL  
 STTLDGDTVLEFDWDPKREIFLHTSAHILAOAVLRIMPAPOTIGPVIDQGFYD  
 FANLISDEDPALTEAMAKTIAEEKPILSROVPEPKALAEFSONPKRAELIAPLE  
 EVELSATYQGEPLDLCRPHLPSTAPVAFKILRTSSAVIKDPSRESILRTYGSFP  
 TTKLEKHLHOLEEAKKRDHRYLTKLDFSOQTSAGAPFHPGMVWVALVDYMK  
 RLHORAGYQOIQTPOLMNRLEMEISGHENKKNNTLVDEEDYAIKPMCGMLY  
 YKTOLHSYREPLRIAEIGHVHRHLSGALSGIMLRPTFHODAHVFLPEVEETL  
 NILNLSYELGFEYHLELSTREOGTIGSDIMELATYKALKRALKVSOKEPTISP  
 GEGAFYGPKIDIHVKDAINRMQCTIOLDMLPRPLKRTNAGCEKSTIMLRAL  
 FGSIRFGLIIEHKGKRPILMSPENRITTVADRHEARAOELKHFSSMGITVSD  
 SSNEVSFKIRNAQMOVNYMTIGDKLETHLLAVRTRDNKVLNDIAVEQFSHILE  
 ELRLSLTPSL"

BASE COUNT 3381 a 2527 c 2095 g 2650 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.51e-57 Length: 10653  
 Score: 1233.50 Matches: 268  
 Percent Similarity: 70.14% Conservatve: 82  
 Best Local Similarity: 53.71% Mismatches: 124  
 Query Match: 52.07% Indels: 25  
 DB: 1 Gaps: 10

US-09-889-314-2 (1-496) x AE001328 (1-10653)

OY 4 MetSerIleSerSerSerGlyProAsp-----AsnGlnLysAsnIleMetSerGln 21  
 DB 4313 AGTCCCTTCATCTTCTTCCTGCTCTCCGATAGCAGCACTTAGATGCTTGTCGAA 4372  
 OY 22 ValLeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGlnThr 41  
 DB 4373 GTCATAGCTTCGACCTCCCAAGCGCTTCTTAATGCAGATAAATTAAACGACATCAGGTT 4432  
 OY 42 LysGlnIleGlnGlnThrArgGlnGlyLys---AsnThrGluMetGluSerAspAlaThr 60  
 DB 4433 AAGCAGTTCACAGCAGCAGCAAAATCGCATGACCTAAGCATGGAAGCGATGTCGCT 4492  
 OY 61 IleAlaMetGlnSerGlyLysAspLysThrSerSerThrThrLys-----ThrGlu 77  
 DB 4493 GTTGCAGGACTGCTGGAAAGATCGCGACGCTTCTGCTCAAAATAGAGACAGCAAGA 4552  
 OY 78 ThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGly 97  
 DB 4553 CTTATAGACAGCAGCAAGATTACTGCGAGGAAGAAGACT-----GCATCT 4597  
 OY 98 AlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLys--- 116  
 DB 4598 GCGGAG-----GCGACATCCCTAAGACCAAGCCGACATTAAGA 4636  
 OY 117 IleAlaMetGlnThrSerIleGlnGluAlaSerLysSerMetGlu---SerThrLeuGlu 135  
 DB 4637 GCTAGCTGCACCAATATGATAGATAGTACTAGCAAACTTTAGAGCTATCTTTTAAGT 4696  
 OY 136 SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaValAlaAla 155  
 DB 4697 TCGTTGTCATCTGTAGATCCACGACATCTACAAAGAAATTCAGATCCCTGCTGCT 4756  
 OY 156 LeuSerGlyLysSerSerLysSer---AlaLysLeuGlnThrProGluLeuProLysPro 174  
 DB 4757 ATGGGTGCTACTACAGAGCTTCTTGACAGCAACTTAGAAACTCCAGACTACCAACCT 4816  
 OY 175 GlyValThrProArgSerGlyValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194  
 DB 4817 TCAAGACACCTCTGCTGTAAGAAGTAATAGCTTGACATTACCAAGCAATTTACC 4876  
 OY 195 ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyraLaserThrGlnAlaGlnAla 214

DB 4877 GCTCTGAGAGCTCAACGACAGCATGTCGAAACCTCCAAAGTACGACATCT 4936  
 OY 215 AspGlnThrAsnLysLeuGlyLysGlnAlaIleLysIleAspLysGluArgGlu 234  
 DB 4937 GCGAATGACCAAAATGCTCTAGATATCTCAAGGCTCTAAATGATTAAGAGCTGAA 4996  
 OY 235 GlnTyGlnGlnMetLysAlaIleGlnGlnLysSerLysAspLeuGlnLysThrMetAsp 254  
 DB 4997 GACTTCAAAAATGCAAGAGATCCAGCAAACTCT---GGAACCAACTCTACATGCAAT 5053  
 OY 255 ThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAla 274  
 DB 5054 ACCGTTAAACAAGTATGATGATGGGTTACCGGCGCTATTACGTGATCTCTAGATATCC 5113  
 OY 275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGlyAla 294  
 DB 5114 GCATTATTCATCTGCGCTGCTGCTGATCGGAACGCTGTCAGAGCCACAGCAGCC 5173  
 OY 295 AlaAlaAlaGlyValAlaAlaGlyAlaAlaAla-----AlaThrThrValAlaThr 311  
 DB 5174 GCGGCTGAGCTACACAGCAGCAGCAGCAGCACTTCTGTAGCTTCAACATGCTGCTACA 5233  
 OY 312 GlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValAlaIleThrAlaValArg 331  
 DB 5234 CAAGTACATGACAGCAGCTGTCGCAAGTGCTTAACACAGCTATATTCAGCTGTAA 5293  
 OY 332 GlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLys 351  
 DB 5294 CAGGCTATGCTCCAAAGCTTATTAACAAGGATTTAAACAAGGATTCACAACAGCATTAAG 5353  
 OY 352 ThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLys 371  
 DB 5354 CAAGCTGTTAAGCGGCTGTGAACCTTCTCTAAACCTGGGTAAATTTTCACGCA 5413  
 OY 372 GlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThr 391  
 DB 5414 GGGAAATATGCTGTATGCAATATGTCCTTAACCTCCAAAGTATTCACAACCTTTGGGA 5473  
 OY 392 SerLysThrValThrValAlaGlyValAlaValAlaAlaProAlaLeuGlyLysGly 411  
 DB 5474 AGTAAATGGGTAACTTACAGATGAGAGCTCTTACAGAGCTTCCAACTGATTCGGG 5533  
 OY 412 IleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPheGlnLysGluVal 431  
 DB 5534 ATTACTAGTCTGAGCTGTGACATGACAGCAAGAAAGAACTGCCCCAAATTCAAAAGAGT 5593  
 OY 432 GlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheProGlnAla 451  
 DB 5594 GGAAGCTCAGACGCTCAATCTGAATGATGAAGAAAGCTTTCACATTTGTCGGCAACAGCA 5653  
 OY 452 SerLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLys 471  
 DB 5654 AGTAAATATGCAAGTAAACAACAGAAAGCCCTTACAGCAACGACAGCGGGCCAA 5713  
 OY 472 LeuGlyAlaGlnIleLeuLysAlaTyraAlaAlaIleSerGlyValAlaIleAlaGlyAla 490  
 DB 5714 ACCGAGCTCAATGACGAAGAGCTTTGCGCAATAGTGGCGCTTACCGCGCGCA 5770

RESULT 38  
 AE002353 9983 bp DNA linear BCT 26-MAY-2000  
 LOCUS Chlamydia muridarum, section 80 of 85 of the complete genome.  
 DEFINITION AE002353 AE002160  
 ACCESSION AE002353.2 GI:8163337  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Chlamydia muridarum.  
 Chlamydia muridarum  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 REFERENCE 1 (bases 1 to 9983) Shen,C., Gill,S.R., Heidelberg,J.F.,  
 White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utteback,T.,  
 Berry,K., Bass,S., Linher,K., Weidman,J., Kouri,H., Craven,B.,



REFERENCE	1 (bases 1 to 10653)	Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R., Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V. and Davis,R.W.
AUTHORS		Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
JOURNAL	Science 282 (5389), 754-759 (1998)	
MEDLINE	9900809	
REFERENCE	2 (bases 1 to 10653)	9784136
AUTHORS		Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Hyman,R.W., Olinger,L., Grimwood,J., Davis,R.W. and Stephens,R.S.
JOURNAL	Comparative genomes of Chlamydia pneumoniae and C. trachomatis	
MEDLINE	9920606	
REFERENCE	3 (bases 1 to 10653)	10192388
AUTHORS		Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R., Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V. and Davis,R.W.
TITLE	Direct Submission	
JOURNAL	Submitted (20-MAY-1998) Program in Infectious Diseases, University of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA	
FEATURES	Location/Qualifiers	
source	1..10653	/organism="Chlamydia trachomatis"
		/strain="D/UW-3/CX"
		/db_xref="taxon:813"
		complement(96..1166)
gene	/gene="pepp"	/note="CT574"
		complement(96..1166)
CDS	/codon_start=1	/protein_id="A6C68176.1"
	/transl_table=11	/product="Amino peptidase P"
	/db_xref="GI:3329016"	
	/translation="MNQNP I K R L D S L V E R A L D A F L I E K D E D I S Y F L O D Q A R S G V L I T R D E A V L F S P D K L D Y A R I O D V L A S Y S K S D O E L G S Y I E N T G L K T I G P D E S Y P Y G I A C K R M N S G Y A F S P O S L V A E K L R C Y S T F E E I O K M R A A F I S A G Y D P V L A I R P R T C E K E L V R M L H Y F W A N L G I E K Y S F P P I A F G N A A P P H A I P I N R S L K G D V L I D I G C Y E G Y C S D M T R I V A F G A T P E Q Q L D L D G Y A V A A G R A A I E F C R A G V P C R D V H K E A V R I L R A H G M E K A F I H G I G H G V G R E H E Y P R L S P F S D A T I Q L M N A V V E P G V Y F P S G G I R I E D T I M I G V E N L N T N K R Y S S E I I I I"	
gene	complement(1177..2907)	
	/gene="mutL"	
	/note="CT575"	
	complement(1177..2907)	
CDS	/codon_start=1	/protein_id="A6C68177.1"
	/transl_table=11	/product="DNA mismatch Repair"
	/db_xref="GI:3329017"	
	/translation="MSLSPSRIRLDSVTANOISAGEVLIENASVYKEIENSDAGA D E I H I E L I G G R G O I V R D N G Y M D E E V P A L O R H A T S K I A H F A D I F S I A S Y G R G E A L P I S A I S I K M E I H T A R A G I G S K T I L E K G E P V C E P A R O G G T I A V H S L F Y N P M R O S P O K S Q M R L A I R R L L N S Y L S S E G I C T M T I S E R O E L Y A K K O G I T E R A L V I G E S P V E A F I D K O O G D L R V L G F L G S P M O H S T O G O R L P I N N R A V E S F T S K K A A V A W M I P A Q R Y P L E V L K T L F P M W C D F N V H P Q T E V R L O E G O I S N L V E A I S P A L R S P S L E E T Y L K Y T E K I P I E N G I S Y P S I R P A I V S A P S C P T S Q P Y L K T E M A T I Y S R D S A S S L Y V E R F R E L T S I G V L L V E D S E G V A V F V O A R K H L F Y S I L S E P R L S R A C O T P L L P S Y O K T K L E A D F L O M R L E A L T A I G I E L S R I S D S P F A I E S A P P I O E B L E K E M I V A L O E B A L H V G E S F E Q L V E N T Y O K L V F S R N A R A P D Y A L D I L M K L G R P E K A F D E M I R R V L D D E M"	
gene	3202..3900	/gene="1crH_1"
	/note="CT576"	
CDS	3202..3900	/gene="1crH_1"
	/codon_start=1	/transl_table=11
	/product="low Calcium Response Protein H"	
	/db_xref="GI:3329018"	
	/translation="MSTPSSNNKSPSPASFNNKSRSLAEIAAQKAKAEDLROKYPVP PTEEFKROYMLDILGSLNGTLIOGLSDVLEIEITVAIVTEFYSOGKVOEALGRLP ILFASKRPOCDYITLISCSYHOLKMYDEAFCFFLAPOAENPPIPPYIADSLMKTIN OPESSODFDYITIDMKCKNPEKYVLKDKCSIMKQSLDVLKKEKSAGKSEFOASPPKKN TKAKKASNNKKAK"	
gene	3917..4276	/gene="CT577"
	3917..4276	/gene="CT577"
CDS	3917..4276	/codon_start=1
	/transl_table=11	/product="hypothetical protein"
	/protein_id="A6C68179.1"	
	/db_xref="GI:3329019"	
	/translation="MSKKRKHQAHTSSKPKYEPAYVSKKSPALQELQNNMISFSD LPIAQMFSETIDCKOLAKMMAALSGMLDSLPETILTKGVFNPKPEAQLSEIISIFL GIKHLTEYVNHKLHADER"	
gene	4313..5776	/gene="CT578"
	4313..5776	/gene="CT578"
CDS	4313..5776	/codon_start=1
	/transl_table=11	/product="hypothetical protein"
	/protein_id="A6C68180.1"	
	/db_xref="GI:3329020"	
	/translation="MSLSSSSSDSSNKNKVIISOYIASTPGCVPMADKLTIDNOYKVO QTRQNRRDLSMESDVAVAGTAKDRRAASQIEGELIEQGLAAGETASADTSITLT QSAKSGASSQOCIEDTSKSLSELSSLSVDATHLOIOSIVSAMQANKELSLTMN ETPGLPKPSTTPROEYMEISLAKAITALGSETQALAEFNOSTOSANNNKSLSES OGLKIDKREEFKMOEIOQKSGTSMVDYKVMIGTVAAITVIVSVALFTGGLL IGRNAAQATAAAGAAATAATTAATVATVATQYNOQYAVOYVAKAIIQYAOIYVOAI KQGITKQIKOAIKOAVKAAVKTLLANKGKITSAGNNAYSKSPKLSKYINTLSKWT LSGVALTAVPOLVSGITSLQSDMOKEIAOIQKEVGALTVAOSEMMKAFITFWDQASK I AAQKTESPSETDQQAKTGAQIATKALASICALAAA"	
gene	5807..7126	/gene="CT579"
	5807..7126	/gene="CT579"
CDS	5807..7126	/codon_start=1
	/transl_table=11	/product="hypothetical protein"
	/protein_id="A6C68181.1"	
	/db_xref="GI:3329021"	
	/translation="MTTGVGNADPDSILAOITONANSASASTGKNQYVAKQOEN VDAFEDLLDQAGTGSGKATANOITSKSGEKAQASSGTTSTVAQAQRTAAQA VHGARDSGFNSDGSATLPSPTGTENVGVLRKGMGLTALMGLIMTLIAQASAKSMSSS FOOONAOINOVAMAPEIGNAIRTOANHOAOATELOAQOSLISGINTLVGAIVSGGG IISASKSIGIKSAATNFTASATISATISAKTATSAIDVDAGTATVAVAKATSGAA SAAKSPAATKTQNMASAKTISQTSASAGLFGQALINTPSWSEKRSRGNNVYKTOG TTRAKFAGRLSSAANI SQWVHGLTAGIDGIVGVIGAVQAOEPMGMAMARAEELK SLNSVOAYASQAQOQLOEQSOOSFNSALQTLQSIDSALQTTASMFN"	
gene	complement(7204..8187)	/gene="CT580"
	complement(7204..8187)	/gene="CT580"
CDS	complement(7204..8187)	/codon_start=1
	/transl_table=11	/product="hypothetical protein"
	/protein_id="A6C68182.1"	
	/db_xref="GI:3329022"	
	/translation="MVAVALLPACTLOTCVFKPCCDMAITFLPLNAFIMSSSALS K SAMEAAPLVEIGSRNVLAVGVLFGLICKRESLTPPPAIMPITVLVSIGFYLTNLV EFITGLQSSSTACFIYGFSPPTAAFCGYVQVLRVYWKKGSLGIVSLVYLILGEG GSEDAVEMQMOLGPELILITATCLSSGYMTLIGRCRSLSMWTAINNAYVIAQVY LSLIHSAVTEVNNPVVENPDLFLIOAGIVTFEMLICVNLIFAKLISFSSFTLSPCN LYNMPLFASFEGWILLGESPPLGLIFAVGNVLGCLNLTITHEEFGCYLTSE"	
gene	8492..10399	/gene="thrS"
	8492..10399	/gene="thrS"
CDS	8492..10399	/codon_start=1
	/transl_table=11	



OY	144	GlnMetLysGluValGluAlaValValValAlaAlaLeuSerGlyLysSerSerIySer	163
Db	904	CAAAAGAAAAGAAGTGGAAAGCGGGTTGTGTCTGCCCTCCACAGGAAAAATTCCGGGGTTCC	963
OY	164	AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle	183
Db	964	GCAAAATTTGGAAACACTGAGCTGCCCAAAGCCCCGGGGTGACACCACAAGATCAGAGGTTATC	1023
OY	184	GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla	203
Db	1024	GAATTCGCAGCTCGCCCTGCTTAAGCAATTCAGACATTGGGAGAACCCACAAATCTGCC	1083
OY	204	LeuSerAsnTyrrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu	223
Db	1084	TTAATCTACTATGACAAGTACACAMACACACAGACCAAAATAAATTAAGTCTAGAA	1143
OY	224	LysGlnAlaIleLysIleAspLysGluArgGluGluTyrrGlnGluMetLysAlaGlu	243
Db	1144	AAGCAAGCATTAATAATCGATTAAAGACAGACAGAAATACCAAGATGAAGGCTCCGAA	1203
OY	244	GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAla	262
Db	1204	CAGAAGTCTAAAGATCTCGAAGCAACATGGAATCTGTCATATCTGATGTCGC	1260
RESULT	36		
LOCUS	E12539	1296 bp DNA linear PAT 27-APR-1998	
DEFINITION		DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.	
ACCESSION	E12539		
VERSION	E12539.1 GI:3251372		
KEYWORDS	JP 1997015243-A/2.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1296)		
AUTHORS	Izutsu,H., Obara,K. and Matsumoto,A.		
TITLE	MEASURING METHOD FOR ANTI-CHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION		
JOURNAL	Patent: JP 1997015243-A 2 17-JAN-1997;		
	HITACHI CHEM CO LTD		
COMMENT	OS None		
	OC Artificial sequences.		
	PN JP 1997015243-A/2		
	PD 17-JAN-1997		
	PF 15-MAR-1996 JP 1996058609		
	PR 28-APR-1995 JP 95P 106009		
	PI IZUTSU HIROSHI, OBARA KAORUHIKO, MATSUMOTO AKIRA PC G01N33/569,C12N1/21,C12N15/02,C12N15/09,C12P21/02,C12P21/08,PC G01N33/53,		
	PC C07K16/12,C07K19/00,		
	C07K16/12.C07K19/00.		
	PC (C12N1/21,C12R1:19),(C12N15/09,C12N1:01),(C12P21/02,C12R1:19);		
	CC Strandedness: Double;		
	CC topology: linear;		
FH	Key	Location/Qualifiers	
FH			
FT	source	1..1296	
FT	mat.peptide	1..1296	
FT		/organism='artificial sequences'	FT
FT		/product='fusion protein of DHFR and antigen peptide'	
FT	misc.feature	1..480	
FT		/note='DHFR coding region'	
FT	misc.feature	484..1296	
FT		/note='antigen peptide coding region'.	
FEATURES	source	Location/Qualifiers	
		1..1296	
		/organism='unidentified'	
		/db_xref='taxon:32644'	
BASE COUNT	407 a 299 c 325 g 265 t		
ORIGIN			

Alignment Scores:		
Pred. No.:	2,866-59	Length: 1296
Score:	1245.00	Matches: 259
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	52.55%	Indels: 0
DB:	6	Gaps: 0
US-09-889-314-2 (1-496) x E12539 (1-1296)		
OY	4 MetSerIleSerSerSerSerGlyProaspasnGlnLysAsnIleMetSerGlnValLeu	23
DB	484 ATGTCTATTTCATCTCTTCTCAGAGACCTACATCAATAAAAAATATCATCTCTCAAGTTCTG	543
OY	24 ThrSerThrProGlnGlnLysValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln	43
DB	544 ACATCGACACCCCCAGGGGGTGGCCCAACAMATTAAGCTGTGTGCCAACGAACCAAGCAA	603
OY	44 IlegInGlnThrArgGlnGlnLysasnThrGlnMetGlnSerAspAlaThrIleAlaGly	63
DB	604 ATACGACCAACCGTCAGAGGTAAACCACTAGATGGAACCGATGCCATATTGCTGTGT	663
OY	64 AlaSerGlyLysAspLysFthrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly	83
DB	664 GCTTGTGGAAAAAGCAAAATCTTCCGACATCAAAAAACAGAAACAGCTCCCAACAGGGA	723
OY	84 ValAlaAlaGlyLysGlnLysSerSerGlnLysAlaGlyAlaAspThrGlyValSer	103
DB	724 GTTGTGTGTGGAAAAAATCCTCAGAAAAAGTCAGAAAGGAGGTGTGATACGTGAGATCA	783
OY	104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle	123
DB	784 GGAGCGGCTGCTACTACAGCATCAATATCTGCAACAAAAATTGCTATGCAACCTCTATT	843
OY	124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnLysLeuGlnSerLeuSerAlaAla	143
DB	844 GAAGAGCGGAGCAAAAGATGAGAGCTACTTAAAGCTCATTCAAAAGCCTCAGTGGCCG	903
OY	144 GlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer	163
DB	904 CAAAGGAAAGAAAGTCGAAGCGGTTGTTGTTCTGCTCCTCAGGGAAGAAAGTTGGGGTTCC	963
OY	164 AlaLysLeuGlnLysThrProGlnLeuProLysProGlyValAlaThrProArgSerGlyValIle	183
DB	964 GCMAAATTGGAAACACCTCGAGCTCCCAAGCCGGGGGTGAACCAAGATTCAGAGTTATC	1023
OY	184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyGlnAlaThrLysSerAla	203
DB	1024 GAAATCGGACATCGCGCTTGTGAAGCATTTAGACATTTGGAGAAAGCCACAAAATTTGCC	1083
OY	204 LeuSerAsnLysThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln	223
DB	1084 TTATTTAACTATGCAAGATACCAAGCACACACAGCCAAACAAATTAACCTVAGGTTCTGGAA	1143
OY	224 LysGlnAlaIleLysIleAspLysGlnLysGlnGlnLysThrGlnGlnMetLysAlaAlaGln	243
DB	1144 AAGCAAGCGATTAATAATGATTAAGACAGACAGATATCCAAAGATGAAGGCTGCCGA	1203
OY	244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla	262
DB	1204 CAGAGTCTAAAGATCTCGAAGGAACATGATGACTGTCACTACTAGTATGATATGCGC	1260
RESULT 37		
LOCUS	AE001328	10653 bp DNA linear BCT 30-Oct-2000
DEFINITION	Chlamydia trachomatis section 55 of 87 of the complete genome.	
ACCESSION	AE001328	AE001273
VERSION	AE001328.1	GI:3329015
KEYWORDS		
SOURCE	Chlamydia trachomatis.	
ORGANISM	Chlamydia trachomatis	
	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.	

US-09-889-314-2 (1-496) x AR122115 (1-1296)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
 DB 484 ATGCTATTTCATCTCTTCTAGAGCTGACAAATCAAAAATATCATGCTCAAGTTCTGT 543

QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnIleThrLysGln 43  
 DB 544 ACATCGACACCCAGGCGGTGCCCCAACAGATAGCTGTCTGGCAACGAAACCAACAA 603

QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
 DB 604 ATACAGCAAAACAGCTGAGGTAAACACAGAGATGGAAGCCATGCCACATTTGCGGT 663

QY 64 AlaSerGlyLysAspLysThrSerSerThrLysThrGlnIleAlaProGlnGlnGly 83  
 DB 664 GCTTCTGAAAAGACAAATCTCTCGACTACAAAACAGAACAGCTCCACAAACAGGA 723

QY 84 ValAlaIleGlyLysGluSerSerSerGlnLysAlaGlyAlaAspThrGlnValSer 103  
 DB 724 GTTGTGCTGCGGAAAGATCTCTCAAAAGTCAAAAGCGAGTGTGATCTGAGATATCA 783

QY 104 G1ValAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 DB 784 GGAACGGCTGCTACTACAGCATCAATATCTGCAACAAAATGCTATGCAACCTCTATT 843

QY 124 GluGlnAlaSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerLeuSerAlaAla 143  
 DB 844 GAAGAGCGCGAACAAAGTATGAGTACCTTACAGTCACTTCAAAACCTCAGTGCAGCG 903

QY 144 GluMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163  
 DB 904 CAATGTAAAGAGCGAAGCGGTGTTGCTGCTCCCTCAAGGAAAGTTCGGGTCC 963

QY 164 AlaLysLeuGlnLysThrProGlnLysProLysProGlyValThrProArgSerGlnValIle 183  
 DB 964 GCAAAATTTGGAACACCTGAGCTCCCAAGCGCGGTGACACCAAGATCAGAGTTATC 1023

QY 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203  
 DB 1024 GAATTCGAGCTCGCGCTTGTCTTAAAGCAATTCAGACATTTGGAGAGACCAAAATCTGCC 1083

QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnIleGln 223  
 DB 1084 TTATCTACTATGCAAGTACACAGCAACAGCAACCAAAATTAATCAATGATCTGAAA 1143

QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysThrGlnGlnMetLysAlaAlaGln 243  
 DB 1144 AAGCAACCGATTAATAATCGATTAAGAACGAGAACATACCAAGATGAAGCTGCGGAA 1203

QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262  
 DB 1204 CAGAAGTCTTAAGATCTCGAAGAACATGATGATCTGATGATGATGCGG 1260

RESULT 35

LOCUS E12526 1296 bp DNA linear PAT 27-APR-1998

DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.

ACCESSION E12526

VERSION E12526.1 GI:3251359

KEYWORDS JP 1997009976-A/2.

SOURCE unidentifed.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1296)

AUTHORS Iizutsu H., Obara K. and Matsumoto A.

TITLE DIHYDROFOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE ANTIBODY

JOURNAL Patent: JP 1997009976-A 2 14-JAN-1997;

COMMENT HITACHI CHEM CO LTD

OS None

CC Artificial sequences.

PN JP 1997009976-A/2

PD 14-JAN-1997

PR 14-MAR-1996 JP 1996057409

PR 28-APR-1995 JP 95P 106006

PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC

C12N15/09, C07H21/04, C07K14/295, C07K14/47, C07K19/00, C12N1/21, PC

C12P21/02,

PC G01N33/569, G01N33/571//A61K39/395, A61K49/00, C12P21/08, PC

(C12N15/09, C12R1:01),

PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91),

PC (C12P21/08,

PC C12R1:91),

CC strandedness: Double;

CC topology: Linear;

FH Key

FH Location/Qualifiers

FT source 1..1296

FT mat\_peptide 1..1296

FT /product='fusion protein of DHFR and antigen peptide'

FT /note='DHFR coding region'

FT misc\_feature 484..1296

FT /note='antigen peptide coding region'

FEATURES

source 1..1296

/organism='unidentified'

/db\_xref='taxon:32644'

BASE COUNT 407 a 299 c 325 g 265 t

ORIGIN

Alignment Scores:

Pred. No.: 2.86e-59 Length: 1296

Score: 1245.00 Matches: 259

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 52.55% Indels: 0

Gaps: 0

US-09-889-314-2 (1-496) x E12526 (1-1296)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
 DB 484 ATGCTATTTCATCTCTTCTAGAGCTGACAAATCAAAAATATCATGCTCAAGTTCTGT 543

QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnIleThrLysGln 43  
 DB 544 ACATCGACACCCAGGCGGTGCCCCAACAGATAGCTGTCTGGCAACGAAACCAACAA 603

QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
 DB 604 ATACAGCAAAACAGCTGAGGTAAACACAGAGATGGAAGCCATGCCACATTTGCGGT 663

QY 64 AlaSerGlyLysAspLysThrSerSerThrLysThrGlnIleAlaProGlnGlnGly 83  
 DB 664 GCTTCTGAAAAGACAAATCTCTCGACTACAAAACAGAACAGCTCCACAAACAGGA 723

QY 84 ValAlaIleGlyLysGluSerSerSerGlnLysAlaGlyAlaAspThrGlnValSer 103  
 DB 724 GTTGTGCTGCGGAAAGATCTCTCAAAAGTCAAAAGCGAGTGTGATCTGAGATATCA 783

QY 104 G1ValAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 DB 784 GGAACGGCTGCTACTACAGCATCAATATCTGCAACAAAATGCTATGCAACCTCTATT 843

QY 124 GluGlnAlaSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerLeuSerAlaAla 143  
 DB 844 GAAGAGCGCGAACAAAGTATGAGTACCTTACAGTCACTTCAAAACCTCAGTGCAGCG 903

OY 184 GIUIEGILEUALALEUALALYSALAIIEGINTHrLEUGLYGUALATHrYSSERALA 203  
|||||  
DB 541 GAAATCGAGACTGCGCTTGCTTAACCAATTCAGACATTGGGAAAGCCCAAAATCTGCC 600  
OY 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLEUGLYGULU 223  
601 TTATCTACTATGCAAGTAGACACAGACAGACCAAAATTAAGTACTAGCTTAGAA 660  
OY 224 LysGlnAlaIleLysIleAspLysGlnArgGluIuTYrGlnGlnLysAlaIleAGLU 243  
661 AAGCAAGCGATTAATAATGATTAAGACAGAGAAATACCAAGATGAAGCTGCCGAA 720  
OY 244 GlnLysSerLysAspLEUGLYGULYThrMetAspThrValAsnThrValMetIleALA 262  
DB 721 CAGAAGCTTAAGAATCTCGAAGGAACAATGATGATCTGTAATCTGTGATGATCGCG 777

RESULT 33  
E12542 813 bp DNA linear PAT 27-APR-1998  
LOCUS E12542  
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.  
ACCESSION E12542  
VERSION E12542.1 GI:3251375  
KEYWORDS JP 1997015244-A/2.  
SOURCE unidentifed.  
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 813)  
AUTHORS Izutsu, H. and Matsumoto, A.  
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION  
JOURNAL Patent: JP 1997015244-A 2 17-JAN-1997;  
HITACHI CHEM CO LTD

COMMENT OS Chlamydia pneumoniae  
PN JP 1997015244-A/2  
PD 17-JAN-1997  
PF 14-MAR-1996 JP 1996057410  
PR 28-APR-1995 JP 95P 106011  
PI IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC G01N33/571, C07K14/295, C07K16/12, G01N33/53,  
G01N33/569//A61K49/00, C12N15/09;  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1. 813  
FT mat\_peptide 1. 813  
FT FT /organism='Chlamydia pneumoniae' strain='YK41',  
/product='antigen peptide'.  
FEATURES  
SOURCE 1. 813  
location/qualifiers  
1. 813  
/organism="unidentifed"  
/db\_xref="taxon:32644"

BASE COUNT 286 a 181 c 187 g 159 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.63e-59 Length: 813  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.55% Indels: 0  
DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12542 (1-813)

OY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
|||||  
DB 1 ATGTCTATTTCATCTCTTCACGAGACCTGACCAATCAAAAAATATCATGTCTCAAGTTCTG 60  
OY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrYsGln 43  
|||||

DB 61 ACATCGACACCCAGGCGGTGCCCAACAGATTAAGCTGTCTGGCAAGAAAGCAAA 120  
OY 44 IIEGInGInThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
DB 121 ATACAGCAAAACACGTCAGAGGTAAACACCTGAGATGGAAGGAGATCCATATTCCTGCT 180  
OY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThAlAProGlnGlnGly 83  
DB 181 GCTTCTGCAAAAGACAAAACCTTCCTCGACTACAAAAACGAAACAGCTCCACAAAGGGA 240  
OY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
DB 241 GTTGCTGCTGGGAAAGAAATCCCAAGAAAGTCAAAAGCAGGTGCTGATACAGACTAGCA 300  
OY 104 GYAlaAlaAlaIleThrThAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
DB 301 GGAGCGGCTGCTACACACATCAATCTGCAACAAAATTTGCTATGACAGCTCTAT 360  
OY 124 GlnGluAlaSerLysSerMetGluSerThrIleGlnGluSerLysLeuSerIleAla 143  
DB 361 GAAGAGCGCAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCTCAGTCCGCG 420  
OY 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaIleSerGlyLysSerSerGlySer 163  
DB 421 CAAATGAAGAAGTGCAGAGCGGTGTGTGCTGCTGCCCTCTCAGGGAAGTTCCGGTTCC 480  
OY 164 AlaLysLysGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183  
DB 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCGGGGTGACACCAAGATCAAGAGTTATC 540  
OY 184 GIUIEGILEUALALEUALALYSALAIIEGINTHrLEUGLYGUALATHrYSSERALA 203  
DB 541 GAAATCGAGACTGCGCTTGCTTAACCAATTCAGACATTGGGAAAGCCCAAAATCTGCC 600  
OY 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLEUGLYGULU 223  
DB 601 TTATCTACTATGCAAGTAGACACAGACAGACCAAAATTAAGTACTAGCTTAGAA 660  
OY 224 LysGlnAlaIleLysIleAspLysGlnArgGluIuTYrGlnGlnLysAlaIleAGLU 243  
DB 661 AAGCAAGCGATTAATAATGATTAAGACAGAGAAATACCAAGATGAAGCTGCCGAA 720  
OY 244 GlnLysSerLysAspLEUGLYGULYThrMetAspThrValAsnThrValMetIleALA 262  
DB 721 CAGAAGCTTAAGAATCTCGAAGGAACAATGATGATCTGTAATCTGTGATGATCGCG 777

RESULT 34  
LOCUS ARI122115 1296 bp DNA linear PAT 16-MAY-2001  
ARI122115  
DEFINITION Sequence 18 from patent US 6165478.  
ACCESSION ARI122115  
VERSION ARI122115.1 GI:14106432  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1296)  
AUTHORS Izutsu, H., Obara, K. and Matsumoto, A.  
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
JOURNAL Patent: US 6165478-A 18 26-DEC-2000;  
FEATURES  
SOURCE 1. 1296  
location/qualifiers  
1. 1296  
/organism="unknown"

BASE COUNT 407 a 299 c 325 g 265 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.86e-59 Length: 1296  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.55% Indels: 0  
DB: 6 Gaps: 0

Query Match: 52.55% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-889-314-2 (1-496) x AR122106 (1-813)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
 1 ATGCTATTTCATCTTCTTCAGACCTGACAAATCAAAAATATCATGCTCAAGTCTCG 60  
 QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 61 ACATGCACACCCACGCGCGCCCAACAAGATAGCTGCTGCAACGAAACAAACAA 120  
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
 121 ATACACCAACACGCTACAGGCTAATAACACTGACATGGAAGCCATGCTATTGCTGCT 180  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
 181 GCTTGGAAAAAGCAAAACTTCTCGACTACAAAAACAAACAGCTCCACAAACAGGA 240  
 QY 84 ValAlaAlaGlySerGlySerSerGlySerGlnLysAlaGlyAlaAspThrGlyValSer 103  
 241 GTTGTGCTGGGGAAGATCTCAGAAAGTCAAAAGCAGCTGCTGATCTGAGATATCA 300  
 QY 104 GlyAlaAlaAlaThrThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 301 GGAGCGGCTGCTACTACAGATCAAAATACGCAACAAAAATTCCTATGACAGCTCTATT 360  
 QY 124 GlnGlnAlaSerLysSerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 361 GAAGAGCGCGCAAAAGTATGAGAGTCTACTTACAGTCAATTCAGAGCCGACAGCGCG 420  
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerGlySer 163  
 421 CAAATGAAAGAACTCGAAGCGGTGTGTGCTGCCCTCTCAGAGAAAGTTGCGGTTC 480  
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183  
 481 GCAAAATTTGGAAACCTGAGCTCCCAAGCCCGGGGTGACCCACCAAGATACAGGTTATC 540  
 QY 184 GlnIleGlyLeuAlaLeuAlaValAlaIleGlnThrLeuGlnGlnAlaThrLysSerAla 203  
 541 GAATGTGAGACTCGCGCTGTGTAAGCAATTCAGACATTTGGAGAGGACCAAAATCTGCC 600  
 QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGln 223  
 601 TTATCTAACTATGCAAGTACACAGCAGACAGCAACCAATAAATACTAGGTCTAGAA 660  
 QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGln 243  
 661 AACCAAGCGTAAATAATGATTAAGACGAGAGCAATACCAAGAGATGAGAGGCTGCCGA 720  
 QY 244 GlnLysSerLysAspLeuGlnGlnThrMetAspThrValAsnThrValMetIleAla 262  
 721 CAGAAAGCTTAAGATCTCGAAGGAACAATGATCTGATCTGATGATGATGCGG 777

RESULT 32  
 E12521 813 bp DNA linear PAT 27-APR-1998  
 LOCUS E12521  
 DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.  
 ACCESSION E12521  
 VERSION E12521.1 GI:3251354  
 KEYWORDS JP 1997009974-A/2.  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 REFERENCE 1 (bases 1 to 813)  
 AUTHORS Izutsu, H. and Matsumoto, A.  
 TITLE ANTIGEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE PATENT: JP 1997009974-A 2 14-JAN-1997;

COMMENT HITACHI CHEM CO LTD  
 OS Chlamydia pneumoniae  
 PN JP 1997009974-A/2  
 PD 14-JAN-1997  
 PF 20-SEP-1995 JP 1995242095  
 PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI  
 IZUTSU HIROSHI, MATSUMOTO AKIRA  
 PC C12N15/09, C07H21/04, C07K14/295, C07K16/12, C12N1/22, C12N15/02, PC C12P21/08  
 PC G01N33/53, G01N33/571//A61K39/118, A61K39/118, A61K39/118, PC (C12N1/22, C12R1:19),  
 PC (C12P21/08, C12R1:91);  
 CC strandness: Double;  
 CC topology: Linear;  
 FH Location/Qualifiers  
 FT source 1..813  
 FT /strain='YK41', /organism='Chlamydia pneumoniae' FT  
 FT mat\_peptide 1..813  
 FT /product='antigen peptide',  
 FT Location/Qualifiers:  
 source 1..813  
 /organism='unidentified'  
 /db\_xref='taxon:32644'

BASE COUNT 286 a 181 c 187 g 159 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.63e-59 Length: 813  
 Score: 1245.00 Matches: 259  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 52.55% Indels: 0  
 DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12521 (1-813)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
 1 ATGCTATTTCATCTTCTTCAGACCTGACAAATCAAAAATATCATGCTCAAGTCTCG 60  
 QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 61 ACATGCACACCCACGCGCGCCCAACAAGATAGCTGCTGCAACGAAACAGGA 120  
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
 121 ATACACCAACACGCTACAGGCTAATAACACTGACATGGAAGCCATGCTATTGCTGCT 180  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
 181 GCTTGGAAAAAGCAAAACTTCTCGACTACAAAAACAAACAGCTCCACAAACAGGA 240  
 QY 84 ValAlaAlaGlySerGlySerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 241 GTTGTGCTGGGGAAGATCTCAGAAAGTCAAAAGCAGCTGCTGATCTGAGATATCA 300  
 QY 104 GlyAlaAlaAlaThrThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 301 GGAGCGGCTGCTACTACAGATCAAAATACGCAACAAAAATTCCTATGACAGCTCTATT 360  
 QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 361 GAAGAGCGCGCAAAAGTATGAGAGTCTACTTACAGTCAATTCAGAGCCGACAGCGCG 420  
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerGlySer 163  
 421 CAAATGAAAGAACTCGAAGCGGTGTGTGCTGCCCTCTCAGAGAAAGTTGCGGTTC 480  
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183  
 481 GCAAAATTTGGAAACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATACAGAGTTATC 540

Db 481 GCAAAATTTGAAACACTGAGCTCCCAAGCCGGGGTGACACCAAGATCAGAGTTATC 540  
Qy 184 GtUtlleglYleuAlaIleuAlaLysAlaIleGIInThrluGIuAlaIaThrlYsSerla 203  
Db 541 GAAATCGAGCTCGCGCTTGAAGCAATTCACATTTGGGAGAAAGCCCAAAATCTGCC 600  
Qy 204 LeuSerAnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGIu 223  
Db 601 TTATCTAACTATGCAGATACACACACAGACAGCAACAAATAAACTAGCTAGAA 660  
Qy 224 LysGlnAlaIleLysIleAspLysGIuArgGIuTYrGlnGlnMetLysAlaIaGIu 243  
Db 661 AAGCAAGCGATTTAAATGATTAAGACGAGAAAGATACCAAGATGATGAGCTGCCGAA 720  
Qy 244 GlnYsSerLysAspLeuGIuGIuThrMetAspThrValaSerThrValMetIleAla 262  
Db 721 CAGAAGTCTAAAGATCTCGAAGAAACAAATGATACTGTCAATACTGTGATGATCGCG 777  
RESULT 30  
E12543 777 bp DNA linear PAT 27-APR-1998  
LOCUS E12543  
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.  
ACCESSION E12543  
VERSION E12543.1 GI:3251376  
KEYWORDS JP 1997015244-A/3.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 777)  
AUTHORS Izutsu,H. and Matsumoto,A.  
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT  
THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA  
INFECTION  
JOURNAL Patent: JP 1997015244-A 3 17-JAN-1997;  
COMMENT HITACHI CHEM CO LTD  
OS Chlamydia pneumoniae  
PN JP 1997015244-A/3  
PD 17-JAN-1997  
PF 14-MAR-1996 JP 1996057410  
PR 28-APR-1995 JP 95P 106011  
PI IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC GOIN33/571,C07K14/295,C07K16/12,G01N33/53,  
PC GOIN33/569//A61K49/00,C12N15/09;  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FH  
FT source 1..777  
FT /organism='Chlamydia pneumoniae' FT  
FT /strain='YK41' /product='antigen peptide'.  
FT /db\_xref='taxon:32644'  
FEATURES  
source Location/Qualifiers  
1..777  
/organism='unidentified'  
/db\_xref='taxon:32644'  
BASE COUNT 277 a 175 c 176 g 149 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.55e-59 Length: 777  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.53% Indels: 0  
DB: 6 Gaps: 0  
US-09-889-314-2 (1-496) x E12543 (1-777)  
Qy 4 MetSerLleSerSerSerGIyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
Db 1 ATGTCTATTTCAATCTCTCTCAGGACCTGACAAATCAAAAAATATCATGCTCAAGTTCTG 60

Qy 24 ThrSerThrProGlnGIyValProGlnGlnAspLysLeuSerGIyAsnGlnThrlYsGln 43  
Db 61 ACATGACACACCACAGGCGGTGCCCAACAAAGATAGCTGTGGCAACGAACGAACAA 120  
Qy 44 IlegInGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaIaThrlleAla 63  
Db 121 ATACGCAAAACCGTCAGGGTAAACACCTAGAGGAAACGATGACCACTATTTGCTGT 180  
Qy 64 AlaSerGIyLysAspLysThrSerSerThrThrlYsThrGlnThrlAlaProGlnGln 83  
Db 181 GCTTTGGAAAAGACAAACTTCCCTCGACTACAAAAAGAAACGACCTCCAAACAGGA 240  
Qy 84 ValAlaIaGIyLysGIuSerSerGIuSerGIuSerGlnYsAlaGIyAlaAspThrGIyAlaSer 103  
Db 241 GTTGCTGTGGGAAAGAAATCTCAGAAAGTCAAAAGCAGGTGCTGATCTGAGATCA 300  
Qy 104 GIyAlaIaIaIaIaThrThrlAlaSerAsnThrlAlaTrpLysIleAlaMetGlnThrSerIle 123  
Db 301 GGAGCGCTGTACTACACGATCAATACTCGAACAAAAATGCTATGACAGACTTATTT 360  
Qy 124 GlnGlnAlaSerLysSerMetGIuSerThrluGIuSerLeuGIuSerLeuSerAlaIa 143  
Db 361 GAAAGAGCGAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAGCCTCAGTCCCG 420  
Qy 144 GlnMetLysGIuValGIuAlaValAlaIaIaIaLeuSerGIyLysSerSerGIySer 163  
Db 421 CAATGAAAGAAAGTCGAAAGCGGTGTGTGCTGCCCTCAGGAAAAAGTTCCGGTTCC 480  
Qy 164 AlaLysLeuGIuThrProGlnLeuProLysProGIyValaThrProArgSerGIuValIle 183  
Db 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCGGGGTGACACCAAGATCAAGGTTATC 540  
Qy 184 GtUtlleglYleuAlaIleuAlaLysAlaIleGIInThrluGIuAlaIaThrlYsSerla 203  
Db 541 GAAATCGAGCTCGCGCTTCTTAAGCAATTCAGACATTGGAGAAAGCCCAAAATCTGCC 600  
Qy 204 LeuSerAnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGIu 223  
Db 601 TTATCTAACTATGCAGATACACACACAGACACCAACAAATAAACTAGCTTAGAA 660  
Qy 224 LysGlnAlaIleLysIleAspLysGIuArgGIuTYrGlnGlnMetLysAlaIaGIu 243  
Db 661 AAGCAAGCGATTTAAATGATTAAGACGAGAAAGATACCAAGATGATGAGCTGCCGAA 720  
Qy 244 GlnYsSerLysAspLeuGIuGIuThrMetAspThrValaSerThrValMetIleAla 262  
Db 721 CAGAAGTCTAAAGATCTCGAAGAAACAAATGATACTGTCAATACTGTGATGATCGCG 777  
RESULT 31  
AR122106 813 bp DNA linear PAT 16-MAY-2001  
LOCUS AR122106  
DEFINITION Sequence 4 from patent US 6165478.  
ACCESSION AR122106  
VERSION AR122106.1 GI:14106423  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 813)  
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.  
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
JOURNAL Patent: US 6165478-A 4 26-DEC-2000;  
FEATURES  
source Location/Qualifiers  
1..813  
/organism='unknown'  
BASE COUNT 286 a 181 c 187 g 159 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.63e-59 Length: 813  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00%  
Query Match: 52.55%  
DB: 6

Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-889-314-2 (1-496) x AR122107 (1-777)

QY 4 MetSer1IeSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
DB 1 ATGCTATTTTCATCTCTTCAGGACCTGCACATCAAAAAATATCATGTCTCAAGTTCTG 60  
QY 24 ThrSerThrProGlnGlnValProGlnAspLysLeuSerGlyAsnGluThrLysGln 43  
DB 61 ACATCGACACCCAGGCGGCGGCCCAACAGATAGCTGTCTGGCAACGAAACAGCA 120  
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
DB 121 ATACAGCAACACGTCAGGAGTAAAAACATGAGATGAAAGCCGATGCCACTATTTGCTGT 180  
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrAlaProGlnGlnGly 83  
DB 181 GCTTCTGGAAGACAAACTCTCTGCATCAAAAAACAGAAACAGCTCCACCAACAGGGA 240  
QY 84 ValAlaIaGlyLysGluSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
DB 241 CTTCTGCTGGGAAGATCTCTGCAAGATCAAAAGCAGGTGCTGATPACTGAGATATCA 300  
QY 104 GlyAlaAlaIaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
DB 301 GGACGCGCTGCTACTACAGCATCAAAATACGCAACAAAAATGCTATGACACCTCTATT 360  
QY 124 GlnGluIaIaSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerIleAla 143  
DB 361 GAAGAGGGAGCAAAAGATGAGTCTTACCTTAGAGTCACTTCAAGCCTTAGTCCGCG 420  
QY 144 GlnMetLysGluValGluAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163  
DB 421 CAATGGAAGAAGTCGAAGCGGTTGTTGCTGCTGCCCTCAGGGAAGTTCCGGTTCC 480  
QY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183  
DB 481 GCAAAATTTGGAACACCGAGCTCCCAAGCCGCGGGAGACCAACAGATCAGAGTTATTC 540  
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyValAlaThrLysSerAla 203  
DB 541 GAATCGGACTCGCGCTTCTTAAGCAATTCAACATTGGGAGAAAGCACAATAATCTGCC 600  
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln 223  
DB 601 TTATCTACTATGCAAGATACAAAGCAACAGCAGCAACAAATAAATAGTGTCTAGAA 660  
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluGluThrGlnGlnMetLysAlaIaGlu 243  
DB 661 AAGCAAGCGATAAATGATTAAGAAAGAGAAATACCAAGAGATTAAGGCTGCCGA 720  
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262  
DB 721 CAGAAGCTTAAGATCTCGAAGAAACATGATGATCTCACTGATGATGCTCCG 777

## RESULT 29

E12522 777 bp DNA linear PAT 27-APR-1998  
LOCUS E12522  
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.  
ACCESSION E12522  
VERSION E12522.1 GI:3251355  
KEYWORDS JP 1997009974-A/3.  
SOURCE unclassified.  
ORGANISM unclassified.

REFERENCE  
1 (bases 1 to 777)  
AUTHORS Izutsu, H. and Matsumoto, A.  
TITLE ANTIGEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE

JOURNAL Patent: JP 1997009974-A 3 14-JAN-1997;  
HITACHI CHEM CO LTD  
COMMENT OS Chlamydia pneumoniae  
PN JP 1997009974-A/3  
PD 14-JAN-1997  
PE 20-SEP-1995 JP 1995242095  
PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI  
IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC C12N15/09, C07H21/04, C07K14/295, C07K16/12, C12N1/21, C12N15/02,  
PC C12P21/08,  
PC G01N33/53, G01N33/571//A61K39/118, A61K39/118, A61K39/118, PC  
(C12N1/21, C12R1:19),  
PC (C12P21/08, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FH FT source 1..777  
FT FT /strain='YK41', /organism='Chlamydia pneumoniae', FT  
FT FT mat\_peptide 1..777  
FT FT Location/Qualifiers  
source 1..777  
/organism='unidentified'  
/db\_xref='taxon:32644'

FEATURES  
source Location/Qualifiers  
1..777  
/organism='unidentified'  
/db\_xref='taxon:32644'

BASE COUNT 277 a 175 c 176 g 149 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.55e-59 Length: 777  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.55% Indels: 0  
DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12522 (1-777)

QY 4 MetSer1IeSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
DB 1 ATGCTATTTTCATCTCTTCAGGACCTGCACATCAAAAAATATCATGTCTCAAGTTCTG 60  
QY 24 ThrSerThrProGlnGlnValProGlnAspLysLeuSerGlyAsnGluThrLysGln 43  
DB 61 ACATCGACACCCAGGCGGCGGCCCAACAGATAGCTGTCTGGCAACGAAACAGCA 120  
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
DB 121 ATACAGCAACACGTCAGGAGTAAAAACATGAGATGAAAGCCGATGCCACTATTTGCTGT 180  
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrAlaProGlnGlnGly 83  
DB 181 GCTTCTGGAAGACAAACTCTCTGCATCAAAAAACAGAAACAGCTCCACCAACAGGGA 240  
QY 84 ValAlaIaGlyLysGluSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
DB 241 GTTGCTGCTGGGAAGATCTCTGCAAGATCAAAAGCAGGTGCTGATPACTGAGATATCA 300  
QY 104 GlyAlaAlaIaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
DB 301 GGAGCGGCTGCTACTACAGCATCAAAATACGCAACAAAAATGCTATGACAGCTCTATT 360  
QY 124 GlnGluIaIaSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerIleAla 143  
DB 361 GAAGAGGGAGCAAAAGATTAAGATCTTACCTTAGAGTCACTTCAAGCCTTAGTCCGCG 420  
QY 144 GlnMetLysGluValGluAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163  
DB 421 CAATGGAAGAAGTCGAAGCGGTTGTTGCTGCTGCCCTCAGGGAAGTTCCGGTTCC 480  
QY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183

Db 887 GGCTGTAAGAAAGCAGACGATATAAATCGATTAAGAGACAGACAGATACCAAGATGAC 946  
OY 241 ALaAlaGlUGlnLysSerLysAspLeuGlnGLYThrMetAspThrValaSerThrValMet 260  
Db 947 GCGCCGCAACAGAGTCTTAAGATCTCGAAGGAGACATGATCTGTCAATCTGATGATG 1006  
OY 261 lIleAla 262  
Db 1007 ATCGCG 1012  
RESULT 27  
E12265 1048 bp DNA linear PAT 27-APR-1998  
LOCUS E12265  
DEFINITION DNA encoding an antigen of Chlamydia pneumoniae.  
ACCESSION E12265.1 GI:3251099  
KEYWORDS JP 1996304404-A/2.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1048)  
AUTHORS Izutsu,H. and Matsumoto,A.  
TITLE DETECTION/MEASUREMENT AND REAGENT FOR ANTICHLADYDIA/ NEUMONIE  
JOURNAL ANTIBODY AND DIAGNOSIS AGENT OFCHLADYDIA/ NEUMONIE INFECTION  
PATENT: JP 1996304404-A 2 22-NOV-1996;  
HITACHI CHEM CO LTD  
OS Chlamydia pneumoniae  
COMMENT PN JP 1996304404-A/2  
PD 22-NOV-1996  
PI 28-APR-1995 JP 1995106014  
PI IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC G01N33/569,C07K14/295,G01N33/53,G01N33/571//A61K9/00, PC  
C12M15/09,C12P21/02,  
PC (C12P21/02,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC fragment\_type: N-Terminal Fragment;  
FH key Location/Qualifiers  
FH source 1. 1048  
FT /organism='Chlamydia pneumoniae' FT  
FT /strain='YK-41' /clone='53-3S' FT  
FT CDS 236..1012  
FT /product='53kDa antigen of Chlamydia FT  
FT pneumoniae' /partial.  
FEATURES  
source Location/Qualifiers  
1. 1048  
/organism='unidentified'  
/db\_xref='taxon:32644'  
BASE COUNT 360 a 241 c 231 g 216 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,67e-60 Length: 1048  
Score: 1262.00 Matches: 262  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
53 Query Match: 53.27% Indels: 0  
Gaps: 0  
DB: 0  
US-09-889-314-2 (1-496) x E12265 (1-1048)  
OY 1 AspThrAsnMetSerLysSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
Db 227 GATACAAACATGCTATTTCATCTCTTCAGACCTGACATCAATAAATAATCATCTCT 286  
OY 21 GlnValLeuThrSerThrProGlnGlnYValProGlnGlnAspLysLeuSerGlyAsnGln 40  
Db 287 CAAGTTCGACATCGACACCCAGGGCGTGCCTCCCAACAAAGATAGCTGTCTGCAACGAA 346

OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60  
Db 347 ACAGACCAATACAGCAACACCTCAGGGTAAAAACACTAGATGGAAACGATGCCACT 406  
OY 61 lIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80  
Db 407 ATTGCTGTCTCTTCTGGAAAAAGACAAACCTTCTCGACTCAAAAACAGACAGCTCCA 466  
OY 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100  
Db 467 CAACAGGAGTCTCTGCGTGGAAAAAGATCCTCAGAAAGTCAAAAAGCGAGTGTGATCT 526  
OY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
Db 527 GGAATACAGAGCGGCTCTACTACAGCATCAAAATCTCAACAAAAATTGCTATGACG 586  
OY 121 ThrSerLleGlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeu 140  
Db 587 ACCTCATTAAGAGGCGACCAAAAGATGAGTCACTTAGATGATCAATCAAGCCTC 646  
OY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaLysSerGlyLysSer 160  
Db 647 AGTCCGCGCAAAATGAAAGAGTCGAAAGCGGTGTGTCTCCCTCCCTCAGGAAAAAGT 706  
OY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
Db 707 TCGGGTTCGCAAAATGGAAACACCTGAGCTCCCAAGCCGGGGGAGACCAAGATCA 766  
OY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThr 200  
Db 767 GAGTATATCAAAATCGACACCTGCGCTTGCATAAAGCATTCAGACATTTGGAGAACGACA 826  
OY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
Db 827 AAATCTGCCATTACTTAATTCAGATGACACAGCACACCAACCAAAATAAATCTA 886  
OY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnMetLys 240  
Db 887 GGCTGTAAGAAAGCAGACGATTAATAATCGATTAAGACAGAGAAAGATTCAGAGATGAAG 946  
OY 241 ALaAlaGlUGlnLysSerLysAspLeuGlnGLYThrMetAspThrValaSerThrValMet 260  
Db 947 GCGCCGCAACAGAGTCTTAAGATCTCGAAGGAGACATGATGATGATCTGATGATG 1006  
OY 261 lIleAla 262  
Db 1007 ATCGCG 1012  
RESULT 28  
AR122107  
LOCUS AR122107 777 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 7 from patent US 6165478.  
ACCESSION AR122107  
VERSION AR122107.1 GI:14106424  
KEYWORDS  
SOURCE unknown.  
ORGANISM unknown.  
REFERENCE 1 (bases 1 to 777)  
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.  
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
JOURNAL Patent: US 6165478-A 7 26-DEC-2000;  
FEATURES  
source Location/Qualifiers  
1. 777  
/organism='unknown'  
BASE COUNT 277 a 175 c 176 g 149 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,55e-59 Length: 777  
Score: 1245.00 Matches: 259  
Percent Similarity: 100.00% Conservative: 0



PH	source	1. .1048	
FT	/strain='YK-41'	/organism='Chlamydia neumonie'	FT
FT			
FT	5'UTR	/clone='53-3S'	
FT	CDS	1. .235	
FT		236. .1012	
FT		/product='53k antigen peptide from Chlamydia	
FT		neumonie'	
FT		1013. .1048.	
FEATURES			
source	location/Qualifiers		
	1. .1048		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	360 a 241 c 231 g 216 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.67e-60	Length:	1048
Score:	1262.00	Matches:	262
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	53.27%	Indels:	0
DB:	6	gaps:	0
US-09-889-314-2 (1-496) x E11479 (1-1048)			
QY	1	AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSer	20
Db	227	GATACAAACATGTCATTATTCATCTTCACAGACCTGCATCAATCAAAAATAATCATGCTCT	286
QY	21	GlnValLeuThrSerThrProGlnGlyAlaProGlnGlnAspIlyLeuSerGlyAsnGln	40
Db	287	CAAGTTCGACATCGACACCCACAGCGCTGCCCAACAAGATAAGCTGTGGCAACGAA	346
QY	41	ThrIlyssGlnIleGlnGlnThrArgGlnGlyAsnThrGlnMetGlnIlySerAspAlaThr	60
Db	347	ACGAAGCAAAATACGCCAACACGTCAGGGTAAACACCTGAGATGGAAGGATGCCACT	406
QY	61	IleAlaGlyAlaSerGlyIlyAspIlyThrSerSerThrThrIlyThrGlnThrAlaPro	80
Db	407	ATTGCTGGTGGCTTCGGAAAAAGACAAACTTCCTCGACTACAAAACAGAACGCTCCA	466
QY	81	GlnGlnGlyAlaAlaAlaGlyIlySglIlySerSerSerGlnIlySalaGlyAlaAspThr	100
Db	467	CAACAGGAGATTGCTGCTGGAAAGAAATCCTCAAAAGTCAAAAGCGAGCTGATAC	526
QY	101	GlyValIlySerGlyAlaAlaAlaThrThrAlaSerSnnThrAlaThrIlyIleAlaMetGln	120
Db	527	GGAGTATCAGGAGCGCGCTGCTACTACACATCAATCAATCAACAAAAATTGCTATGCAG	586
QY	121	ThrSerIleGlnGlnAlaSerIlySerMetSlnSerThrIleuGlnIlySerIleu	140
Db	587	ACCTTATTGAAGAGCGCACGAAAGTATGGAGTCTACTTAGAGTCACTTCAAAAGCTC	646
QY	141	SerAlaAlaGlnMetIlySglIlyAlaGlnAlaValAlaValAlaAlaLeuSerGlyIlySer	160
Db	647	AGTGCCGGCGCAATGAAGAAGTGCAGAACCGGTTGTTGCTGCCCTCAGGGAAAAAGT	706
QY	161	SerGlySerAlaIlySglIlyThrProGlnIlyuProIlyProGlyValThrProArgSer	180
Db	707	TTCGGTTCCCGCAAAATTTGGAAACACTTGAGCTCCCAAGCCCGGGGATGACCCAAAGTCA	766
QY	181	GlnValIleGlnIleGlyIlyAlaLeuAlaIlySalaIleGlnThrIleuGlnIlyAlaThr	200
Db	767	GAGGTTATCGAAATCGGACTCGCGTCTCTAAGCAATTCAGCAATTGGGAGAACCCACA	826
QY	201	IlySerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnIlyLeu	220
Db	827	AAATCTGCTTATCTAACTATATGCAAGTACACAGACACAGACGACCAACAAATAAACTA	886
QY	221	GlyIleuGlnIlySglIlyAlaIleIlySlnAspIlySgluArgGlnIlyrGlnIlyMetIlyS	240



Oy	64	AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGly	83
Db	720	GCCTTCTGGAAAAGACAAACCTCTCTCGACTACAAAAGAAACAGCTCCACACAGGA	779
Oy	84	ValAlaIaGlyLysGluSerSerGluSerGlnLysAlaGlyValAspThrGlyValSer	103
Db	780	CTTCCTCTCGGAAAAGAACTCTCCAGAAAGTCAAAGCGAGCTGATACGTAGATCA	839
Oy	104	GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle	123
Db	840	GGAGCGGCTGCTACTACAGCATCAAACTACTGCACAAAAATTCATACACACTATT	899
Oy	124	GluGluAlaSerLysSerMetGluSerThrIleGluSerLeuGlnSerLeuSerAla	143
Db	900	GAAAGGGGACCAAACTATGAGAGTCTACCTTAGATCACTTCAAAGGCTCAGTGGCGG	959
Oy	144	GlnMetLysGluValGluAlaValValAlaAlaLeuSerGlyLysSerSerGlySer	163
Db	960	CAAAATGAAAGAGTCGAAGCGGTTGTGTGTGCTCGCCCTCAGGAAAAAGTTCCGGGTCC	1019
Oy	164	AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle	183
Db	1020	GCAAAATTTGAAACCTGTAGCTCCCAAGCCCGGGGTATACCAAGATCAGAGTTATC	1079
Oy	184	GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlyAlaThrLysSerAla	203
Db	1080	CAATTCGAGCTCGCGCTTGGTAAAGCATTCAGACATTGGAGAAAGCCACAAATCTGCC	1139
Oy	204	LeuSerAsnTyraIaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu	223
Db	1140	TTATCTTAACATATGCAGACTACCAAGCACAAGCCAGCCAAACAAATTAACCTAGGCTTAGAA	1199
Oy	224	LysGlnAlaIleLysIleAspLysGluValArgGluGluThrGlnMetLysAlaIaGlu	243
Db	1200	AAGCAGCGATAAAAATCGATTAGAACAGCAGAGAACTACCAAGAGTGAAGGCTGCGAA	1259
Oy	244	GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaVal	263
Db	1260	CAGAAAGCTAAAGATCTCGAAGCAACAATGATACGTCAATATCTGTATGATGCCGAA	1318
Oy	264	SerValAlaIleThrVal-----IleSerIleValAla	274
Db	1319	GGGGTGCAAATGGCAATGGGGGGCCCTTAATTAAATTACTCGAGAGATCCAGATCTTAATCG	1378
Oy	275	AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGlyAlaAlaVal	292
Db	1379	ATGATCCCTTAGCGGCGGAGCCATCTGTGCCGCGATCACCGGGGCCACAGTGCGGTTGCT	1438
Oy	293	GlyAlaAla-----	295
Db	1439	GGCGGCTA-TATCGCGACATCACCGATGGGAAAGATCGGGCTCCCACTTCGGGCTCAT	1497
Oy	296	-----AlaAlaGlyGlyValAlaIaGlyVala	303
Db	1498	GAGCGCTTGTGGCGCTGGGTATGGTGGCAGGCCCGCTGGCCGGGGACTGTGGGCGCC	1557
Oy	304	-----AlaAlaAla-----	306
Db	1558	ATTCCTTGCATGACCATTCCTTGGCGGGCGGTGCTCAAGGGCTCAACCTACTACTG	1617
Oy	306	-----	306
Db	1618	GGCTGCTTCTTAATGCAGAGTGCATAMGGAGAGCGTCGACCGATGCCCTTGAGAGCC	1677
Oy	307	-----ThrThrValAlaThrGlnIle	313
Db	1678	TTCAACCGACATGCACTCTCCGGTGGGGCGGGGCGATGATATATGTGGCCGCAATTATG	1737
Oy	314	ThrValGlnAlaValValGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAla	333
Db	1738	ACTGCTCTTCAATCAGCAACTGTAGGACAG-----	1770

```

QY 334 lIeThrAlaAlaIle-----LysAlaAlaValLys 343
      ||| ||| |||
Db 1771 GGCcCGcAGcGGcCTTGgGcTATTTCGcGAGcAGcACcCGTTTCcCTGcAGcCAGcATG 1830
QY 344 SerGlyIleLysAlaAPheIleYthrIeuValLysAlaIleAlaLysAlaIleSerLys 363
      ||| ||| |||
Db 1831 ATcGGcCTGcTcCTTGcGGcTATTCGGAArCTTGcAGcGcCTTGcTCAAGcCTTGcTACT 1890
QY 364 GlyIleSerLysValAPheAlaLysGlyIthrGlnMetIleAlaLysAsnPheProLysLeu 383
      ||| ||| |||
Db 1891 GcTCcCGcCGc-----ACCAAAcCTTGcGGcGAGcAAG 1920
QY 384 SerLysValIleSerSerLeu----- 390
      ||| ||| |||
Db 1921 CAcGGcATTATTCGcCGcGATGcGGcCGcAGcCGcCTGcCTACGCTTCGcTGGcCTTCGCG 1980
QY 391 ThrSerLysTrpValrIhrValGlyValGlyValAlaAlaAlaProAlaLeuGlyLys 410
      ||| ||| |||
Db 1981 AcGGcAGcGTGATGcCGcCTTCcCCcCTTATGATrCTTCrCGcCTTGcGGcGcGATGcGGcATG 2040
QY 411 GlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGln 430
      ||| ||| |||
Db 2041 cCGcGGcTTCAGcAGcCATGcGTGcTCrCAGcGcAGcGTAGATGcAGcACcATCAGcGcAGcCTTCA 2100
QY 431 ValGlyLysLeuGlnAlaAlaAlaAlaAspMetIleSerMet 443
      ||| ||| |||
Db 2101 ---GGATcGCTTCGcCGcCTTCrACcAGcCTTACcCTTACCTTGcATG 2136

```

RESULT 25	LOCUS	AR122109	1048 bp	DNA	Linear	PAT 16-MAY-2001
DEFINITION	Sequence	AR122109	9 from patent US 6165478.			
ACCESSION	AR122109					
VERSION	AR122109.1	GI:14106426				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 1048)					
TITLE	Izutsu, H., Obara, K. and Matsumoto, A.					
JOURNAL	DNA encoding Chlamydia pneumoniae antigenic polypeptide					
FEATURES	Patent: US 6165478-A 9 26-DEC-2000;					
	Location/Qualifiers					
	source	1..1048				
		/organism="unknown"				
BASE COUNT	360 a	241 c	231 g	216 t		
ORIGIN						
Alignment Scores:						
Pred. No.:	2.67e-60	Length:	1048			
Score:	1262.00	Matches:	262			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	53.27%	Indels:	0			
DB:	6	Gaps:	0			
US-09-889-314-2 (1-496) x AR122109 (1-1048)						
Qy 1	AspThrAsnMetSerIleSerSerSerSerCylProAspAsnGlnLysAsnIleMetSer 20					
Db 227	GATPACAAACATGTCATATTTCATCTTCTTCAGGACCTGCAACATCAAAAATAATCATGTCT 286					
Qy 21	GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40					
Db 287	CAAGTCTGACATCGACACCCAGGAGCGTCCCAACAAGATMACTGTCTGGCAACGAA 346					
Qy 41	ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThr 60					
Db 347	ACGAACCAATATACAGCAAAACACGTACGCTAATAAAACACTGAATGGAAGCAATGCCACT 406					
Qy 61	IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80					
Db 407	ATTGCTGGTGGCTTCTGGAAAGCAAAACTTCTGACTACAAAAACAGAAACACTCCA 466					

Db 960 CAATGAAGAAGTCGAAGGGTGTGTGTCGCCCTCAGGAGAAAGTTCGGCTCC 1019  
QY 164 AAlaIleuGlInthrProGluLeuProLysProGlyValIthrProArgSerGluValIle 183  
Db 1020 GCAAAATGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGGTTATC 1079  
QY 184 GluIleGlyLeuAlaLeuAlaIleGlyIthrLeuGlyGluIleThrIleSerAla 203  
Db 1080 GAAATCGAGCTCGGCTGTCTTAAAGCAATTCAGACATGGAGAGCCACAAATCTGCC 1139  
QY 204 LeuSerAntIyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223  
Db 1140 TTATCTTAACATGCAATACACAGCAGACAGACCAACAATTAAGCTGCTGAA 1199  
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluIuTyrGlnGluMetLysAlaGlu 243  
Db 1200 AAGCAACGATTAATAATGATTAAGAACAGAAAGATACCAAGATGAAGAGCTCCGAA 1259  
QY 244 GlnIySerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 1260 CAGAAAGCTTAAAGATCTCGAAGAACAAATGATCTGATCTGATGATGCGAA- 1318  
QY 264 SerValAlaIleThrVal-----IleSerIleValAla 274  
Db 1319 GGGTTGCAATGGCATGGGGCCCTTAATTAATTAATCAGAGATCCAGATCTAATCG 1378  
QY 275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGlyAlaAlaVal 292  
Db 1379 ATGATCTCTTAACGGGAGCAGCATGTGGCGGCATCCGCGCAGACAGTGGCGTGTCT 1438  
QY 293 GlyAlaAla----- 295  
Db 1439 GGGCCCTTA-TATGCCGACATCCAGATGGGAAAGATCGGCTCGCCACTTCGGGCTCAT 1497  
QY 296 -----AlaAlaGlyAlaAlaGlyAla 303  
Db 1498 GAGGCTTTGTTGGCGTGGGTATGTGGCGGCGCGGCGGCGGAGCTTTGGGCGCC 1557  
QY 304 -----AlaAlaAla----- 306  
Db 1558 ATCTCTTGCATGCACCATTCCTTCCGCGCGGCTGCTCAAGGCTCAACCTACTACTG 1617  
QY 306 ----- 306  
Db 1618 GCGCTTCTTAATGCAGAGAGTCCATTAAGGAGAGCGTCGACCGATGCCCTTAGAGACC 1677  
QY 307 -----ThrThrValAlaThrGlnIle 313  
Db 1678 TTCAACCCAGTCAGCTCTTCGGTGGGCGGCGGCGCATGACTATTCGCGCGCACTTATG 1737  
QY 314 ThrValGlnAlaValIleGlnAlaValIleLysGlnAlaValIleThrAlaValIleArgGlnAla 333  
Db 1738 ACTGTCTTCTTATCATCATCACTGCTAGAGCAG----- 1770  
QY 334 IleThrAlaAlaIle-----LysAlaAlaValLys 343  
Db 1771 GTGGCGGAGGCTCTGGGTATTTTGGCGAGACCGCTTTCGCTGAGGCGCAGCATG 1830  
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
Db 1831 ATCGGCGCTGCGCTTGGGATTTGGGAATCTTGACGCGCTTCGCTCAAGCTTGCTCACT 1890  
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
Db 1891 GGTCCCGCC-----ACCAGAAAGCTTTCGGCGAGAG 1920  
QY 384 SerLysValIleSerSerLeu----- 390  
Db 1921 CAGGCCATTAATCGCGGATCGGCGAGCGCGCTGGGCTGCTGCTGGCTCGCG 1980  
QY 391 ThrSerLysTrpValIleThrValGlyValGlyValValAlaIleProAlaLeuGlyLys 410  
Db 1981 ACGGGAGGCTGTGATGGCTTCCCATTTATGATTTCTTCGCTTCGGGGCATCGGATG 2040

QY 411 GlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLysGlu 430  
Db 2041 CCCCGTTGGAGCCATGCTGTCCAGGACAGTACATGACATCACTCAGGACAGCTTCAA 2100  
QY 431 ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
Db 2101 ---GGATCGCTCGGCGCTTACCAAGCTTAATCTGCATC 2136  
RESULT 24  
E12540  
LOCUS E12540  
DEFINITION Nucleotide sequence of pcPN533T that contains Chlamydia pneumoniae antigen gene.  
ACCESSION E12540  
VERSION E12540.1 GI:3251373  
KEYWORDS JP 1997015243-A/3.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 5438)  
AUTHORS Iizusu,H., Obara,K. and Matsumoto,A.  
TITLE MEASURING METHOD FOR ANTI-CHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION  
JOURNAL Patent: JP 1997015243-A 3 17-JAN-1997;  
HITACHI CHEM CO LTD  
OS None  
OC Artificial sequences.  
PN JP 1997015243-A/3  
PD 17-JAN-1997  
PR 15-MAR-1996 JP 1996058609  
PR 28-APR-1995 JP 95P 106009  
PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC  
G01N33/569, C12N1/21, C12N15/02, C12P21/02, C12P21/08, PC  
G01N33/53,  
PC G01N33/531, G01N33/571//A61K49/00, C07K14/245, C07K14/39, PC  
C07K16/12, C07K19/00,  
PC (C12N1/21, C12N1.19), (C12N15/09, C12R1.01), (C12P21/02, C12R1.19);  
CC strandedness: Double;  
CC topology: Circular;  
FH Key Location/Qualifiers  
FH FT source 1..5438  
FT FT Location/Qualifiers  
FEATURES  
source 1..5438 /organism='Artificial sequences'.  
BASE COUNT 1372 a 1424 c 1391 g 1251 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.33e-59 Length: 5438  
Score: 1265.00 Matches: 312  
Percent Similarity: 62.39% Conservative: 33  
Best Local Similarity: 56.42% Mismatches: 75  
Query Match: 53.40% Indels: 135  
Gaps: 10  
US-09-889-314-2 (1-496) x E12540 (1-5438)  
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
Db 540 ATGCTATTATTCATCTTCTTCAGGACCTGACATCAAAAAATATCATGCTCAAGTCTCG 599  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLysLeuSerGlyAsnGlyThrLysGln 43  
Db 600 ACATCGACACCCCAAGGGGTGCCCCACACAGATTAAGCTGTGTGGACAGAAACAGACAA 659  
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnLeuSerAspAlaThrIleAlaGly 63  
Db 660 ATACGCAAAACAGTCAGAGTAAAAACACTGAGATGGAAGCAAGCATGCCACTATTGCTGT 719

Oy	264	SerValAlaIleThbVal-----	IlSerIleValAla	274
Db	1319	GGGGTTCGAATTTGCCATGGGGCCCTTAATTATTAACTCGAGAATTCAGATCTTAATCG		1378
Oy	275	AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----	AlaAlaGlyAlaAlaVal	292
Db	1379	ATGATCCTCTACCGCGGACGCAATCGTGGCGGCATCACCGGCGCACAGATCGGGTTGCT		1438
Oy	293	GlyAlaAla-----		295
Db	1439	GAGGCGCTA-TATCGCGCAATCCAGATGGGAGATCGGGCTCGCACTTCGGGCTCAT		1497
Oy	296	-----	AlaAlaGlyAlaAlaVal	303
Db	1498	GAGCGCTTGTTTGGCGGATGGTATGGTGGCAGGCCCGCTGGCGGGGAGACTGTGGGGCGC		1557
Oy	304	-----	AlaAlaAla-----	306
Db	1558	ATTCCTTCGATGACCACTTCCTTCGGCGGCGGTCTCAACGGCCTCAACTACTACG		1617
Oy	306	-----		306
Db	1618	GGCTGCTTCCTTAATCAGAGATGCGCATAGGAGAGCTCGACCGATCCCTTGAGACC		1677
Oy	307	-----	ThrThrValAlaThrIle	313
Db	1678	TTCAACCCAGTCAGCTCCTTCGCGTGGCGCGGCGGCATGACATGCTGCGCCACTTATG		1737
Oy	314	ThrValGlnAlaValAlaGlnAlaValAlaGlnAlaValIleThrAlaValArgGlnAla		333
Db	1738	ACTGCTCTTATATCATCTCACTCTAGACAG-----		1770
Oy	334	IleThrAlaAlaIle-----	LysAlaAlaValLys	343
Db	1771	GTGCCGGGACGGCTCTGGGTCAATTTCCGGCGAGAACCCGTTTCCTCGGAGCCGACGATG		1830
Oy	344	SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys		363
Db	1831	ATCGGCTCGTGGCTTCCGATTCGCAATCTGCAGCCCTCGCTCAAGACCTTCGCTACT		1890
Oy	364	GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu		383
Db	1891	GGTCCGCC-----	ACCAAACTTTCGGCGAGAG	1920
Oy	384	SerLysValIleSerSerLeu-----		390
Db	1921	CAGGCAATTATCGCGGATGGCGGCGACGCGCTGGGTACGTCCTGGCGCTTCGCG		1980
Oy	391	ThrSerLysTPValThrValGlyValGlyAlaValAlaAlaPheProAlaLeuGlyLys		410
Db	1991	ACGCGAGGCTGGATGGCTTCCCTCCATTAATGATCTTCGCTTCGCGGCGCATCGGGATG		2040
Oy	411	GlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGlu		430
Db	2041	CCCCGCTTGCAGAGCCATCGCTGTCAGGCGAGTGAAGACACATCAGGAGACGCTTCA		2100
Oy	431	ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet		443
Db	2101	--GGATGCTCGCGGCTTACACGACCTAATCTTCGATC		2136
RESULT	23			
LOCUS	E12527	5438 bp	DNA	linear
DEFINITION	E12527			PAT 27-APR-1998
ACCESSION	E12527			antigen gene.
VERSION	E12527.1			
KEYWORDS	JP 1997009976-A/3.			
SOURCE	unidentified.			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 5438)			
AUTHORS	Izutsu,H., Obara,K. and Matsumoto,A.			

TITLE	DIHYDROPHOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE ANTIBODY
JOURNAL	Patent: JP 1997009976-A 3 14-JAN-1997; HINACHI CHEM CO LTD
COMMENT	OS None OC Artificial sequences. PN JP 1997009976-A/3 PD 14-JAN-1997 PF 14-MAR-1996 JP 1996057409 PR 28-APR-1995 JP 95P 106006 PI IZUMSU HIROSHI, OHARA KAZUHIKO, MATSUOKO AKIRA PC C12P21/02, C12P21/02, PC G01N33/569, G01N33/571/A61K39/395, A61K49/00, C12P21/08, PC (C12N15/09, C12R1:01), PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91), PC (C12P21/08, PC C12R1:91); CC strandedness: Double; CC topology: Circular; FH Key Location/Qualifiers
FEATURES	FT source 1..5438 FT /organism='Artificial sequences', Location/Qualifiers 1..5438 /organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	1372 a 1424 c 1391 g 1251 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.33e-59
Score:	1265.00
Percent Similarity:	62.39%
Best Local Similarity:	56.42%
Query Match:	53.40%
DB:	6 Gaps: 10
US-09-889-314-2 (1-496) x E12527 (1-5438)	
QY	4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
Db	540 ATGCTATTTCATCTCTTCACGACCTGACATCAATCAAAATATCATCTCTCAAGTTCTG 599
QY	24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGln 43
Db	600 ACATGCACACCCCAAGGCGTCCGCCAACAAATAGCTGTGGCCAAACGAAACGAA 659
QY	44 IlegInGInThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63
Db	660 ATACGCAAAACCGCAGAGGTAAACACTGAGATGAGAAAGCGATGCCATATTGCTGCT 719
QY	64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
Db	720 GCTTCTGGAAGAAAGCAAAACTTCTCCGCACCTACAAAAGACGAACAGCTCCACACAGGA 779
QY	84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db	780 GTTGTCTGCTGGAAAGATCTCTCAGAAAGTCAAAAGCGAGTGCATACGAGATCA 839
QY	104 GlyAlaAlaAlaThrThrAlaSerSerSerThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db	840 GGAGGCGCTGCTACTACAGCATCAATACTGTCACAAATAATTGCTATGCAACCTTATT 899
QY	124 GluGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
Db	900 GAAGGCGGAGCAAAAGATGAGAGCTTACCTTAAGTACTTCAAAAGCTCAGTGCCTGG 959
QY	144 GlnMetLysGluValGluAlaValValAlaAlaLeuSerGlyLysSerSerLysSer 163

QY	241	AlaAla	glu	gln	lys	ser	lys	asp	leu	glu	gln	thr	met	asp	thr	val	met	260				
Db	1471	GCTGCGCAACGAAGTCTAAAGATCTCGAAGAACAAATGATCTGTCATAATCTGTATG																1530				
OY	261	IleAlaVal	ser	Val	Ile	Thr	Val	-----									Ile	ser	271			
Db	1531	ATCCGCGAA-GGGGTTTCGAATTGCCATGGGGGCCCTTAATTATTAATCACTCGAGAAATCCAG																1589				
OY	272	Ile	Val	Ala	Ile	Ala	leu	Thr	Cys	Gly	Val	Ala	Gly	Leu	Ala	Gly	Leu	-----	Ala	Ala	Gly	289
Db	1590	ATCTAATATGAGTAATCCTTACGCGGACGATCGTGGCGGCGATACCGGGCCACAGAGT																				1649
OY	290	AlaAla	Val	Gly	Val	Ile	Ala	-----														295
Db	1650	GCGGTTGCTGGGGCTTA-TATCGCGACATCACGATGGGAGACATCGGGCTCGCCACTT																				1708
OY	296	-----																				300
Db	1709	CGGGCTCATGAGCGCTTGTTCGGCGGTGGATGATGATGTCAGAGCCCGGTGGCGGGGAGCTG																				1768
OY	301	Ala	Gly	Ala	-----																	306
Db	1769	TTGGCGCGCATTCCTCTTCATGACCATTCCTTCTGGCGGGCGGTCTCAACGCGCTTCAC																				1828
OY	306	-----																				306
Db	1829	CTACTACTGGGCTGCTTCTCTTATGACGAGATGCGCATAGGAGAGAGCTCGACCATGCCC																				1888
OY	307	-----																				310
Db	1889	TTGAGAGCCTTCACCCACGTCAGTCTCTCCGTGGCGCGCGGGCGAGTCATGCTGCC																				1948
OY	311	Thr	gln	Ile	Thr	Val	Gln	Ala	Val	Ala	Val	Gln	Ala	Val	Ile	Thr	Ala	Val	-----			330
Db	1949	GCACCTAATGACGTCTTCTTATCATGCACTCGTAGGACAG																				1990
OY	331	Arg	Gln	Ala	Ile	Thr	Ala	Ala	Ile	-----												340
Db	1991	-----																				2041
OY	341	Ala	Val	Lys	ser	Gln	Ile	Lys	Ala	leu	Thr	Leu	Val	Lys	Ala	Ile	Ala	Lys	Ala			360
Db	2042	GCGACGATGATCGGCTGCTGCTCGGTATTCGGAATCTTGACACGCCCTCGCTCAAGCC																				2101
OY	361	Ile	ser	Lys	Gly	Ile	ser	Lys	Val	Ala	leu	Ala	Lys	Gly	Thr	Gln	Met	Ile	Ala	Lys	Asn	380
Db	2102	TTTCGTCACCTGGTCCGCC																				2131
OY	381	Pro	Lys	Leu	ser	Lys	Val	Ile	ser	ser	Leu	-----										390
Db	2132	GCGCGAAGACGCGCATATACCGCGCATGGCGGCGGACGCGCGCTACGCTTGTG																				2191
OY	391	-----																				407
Db	2192	GCGTTCGGGACGCGAGCGGTGATGTCCTTCCCATTAATGATCTCTGCTCCGCGGCGC																				2251
OY	408	Leu	Gly	Lys	Gly	Ile	Met	Gln	Met	Ile	Leu	ser	Gln	Met	Gln	Asn	Val	Ala	Ile	Ala	Ile	427
Db	2252	ATCCGAGATGCCCGCGCTGAGCGCATGCTGACGACGAGCATAGATGACGACCATCAGGGA																				2311
OY	428	Gln	Lys	Gln	Val	Gly	Lys	Leu	Ile	Ala	Ile											

REFERENCE	Unclassified.			
AUTHORS	Izutsu,H., Obara,K. and Matsumoto,A.			
TITLE	DNA encoding Chlamydia pneumoniae antigenic polypeptide			
JOURNAL	Patent: US 6165478-A 25-Dec-2000;			
FEATURES	Location/Qualifiers			
SOURCE	1..5438			
BASE COUNT	1372 a	1424 c	1391 g	1251 t
ORIGIN	/organism="unknown"			
Alignment Scores:				
pred. No.:	1.33e-59	Length:	5438	
Score:	1265.00	Matches:	312	
Percent Similarity:	62.39%	Conservative:	33	
Best Local Similarity:	56.42%	Mismatches:	75	
Query Match:	53.40%	Indels:	135	
DB:	6	Gaps:	10	
US-09-889-314-2 (1-496) x AR122122 (1-5438)				

QY 296 -----AlaAlaGlyGlyAla 300  
DB 1709 GCGGCTCATGAGCGGCTGTTGTCGGGGTGGATGGTGGAGGCCGCTGGCGGGGAGCTG 1768  
QY 301 AlaGlyAla-----AlaAlaAla----- 306  
DB 1769 TTGGGGCGCCATCTCTGTCATGCACCATTCCTTCTGGCGGGCGTGTCAAGCGCTCAAC 1828  
QY 306 ----- 306  
DB 1829 CTACTACTGGGCTGCTCTCTTAATGACAGAGTCCATTAAGGAGAGCGTCGACCATGCC 1888  
QY 307 -----ThrThrValAla 310  
DB 1889 TTGAGAGCCTTCACACCAAGTACGCTCTCGGTGGGGCGGGGACATGACTATGCTGCGCC 1948  
QY 311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330  
DB 1949 GCACCTATGACGCTCTCTCTTATCATGCMACTGACAGACAG----- 1990  
QY 331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340  
DB 1991 -----GTCCGCGACGCGCTGCTGGGTCAATTTGCGGAGACCGCTTGCTGGAGAC 2041  
QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
DB 2042 GCGAGATGATGAGGCTGCTGCGCTGCGGATTCGGATCTTCACAGCCCTGCTCAAGCC 2101  
QY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnIle 380  
DB 2102 TTCTGTCAGTGGTCCGCGC-----ACCAAGCTTTC 2131  
QY 381 ProLysLeuSerLysValIleSerSerLeu----- 390  
DB 2132 GCGGAGACGACGCGCATTTATGCGCGCATGGCGCGCGCGCTGAGCTTCTGCTG 2191  
QY 391 -----ThrSerLysTrpValThrValGlyValGlyValAlaValAlaAlaProAla 407  
DB 2192 GCGTTCGCGACGCGAGCGATGGCTTCCCATTAATGATCTTCTGCTTCCGGCGGC 2251  
QY 408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPhe 427  
DB 2252 ATCGGAGATGCCCGCTGTGCGCGCATGCTGCAGGACGATGATGACCATCAGGGA 2311  
QY 428 GlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
DB 2312 CAGCTTCAA--GGATCGCTCGCGGCTTTACACGCTAATTGCAATC 2356  
RESULT 21  
E12544 5658 bp DNA linear PAT 27-APR-1998  
LOCUS E12544  
DEFINITION Nucleotide sequence of pCPM533 alpha that contains Chlamydia  
pneumoniae antigen gene.  
ACCESSION E12544  
VERSION E12544.1 GI:3251377  
KEYWORDS JP 1997015244-A/4.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 5658)  
AUTHORS Izutsu,H. and Matsunoto,A.  
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT  
THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA  
INFECTION  
PATENT: JP 1997015244-A 4 17-JAN-1997;  
JOURNAL HITACHI CHEM CO LTD  
COMMENT  
OS None  
OC Artificial sequences.  
PN JP 1997015244-A/4  
PD 17-JAN-1997  
PF 14-MAR-1996 JP 1996057410  
PR 28-APR-1995 JP 95P 106011  
PI IZUTSU HIROSHI, MATSUMOTO AKIRA

PC G01N33/571,C07K14/295,C07K16/12,G01N33/53,  
PC G01N33/569//A61K49/00,C12N15/09;  
CC strandedness: Double;  
CC topology: Circular;  
FH key Location/Qualifiers  
FH FT source 1..5658  
FEATURES  
source /organism="Artificial sequences".  
BASE COUNT 1446 a 1477 c 1428 g 1307 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.68e-60 Length: 5658  
Score: 1282.00 Matches: 315  
Percent Similarity: 62.59% Conservative: 33  
Best Local Similarity: 56.65% Mismatches: 75  
Query Match: 54.12% Indels: 135  
DB: Gaps: 10  
US-09-889-314-2 (1-496) x E12544 (1-5658)  
QY 1 AspThrAsnMetSerLysSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
DB 751 GATACAAACATGTCATTTTCATCTTCTTCAGGACCTGCACAAATCAAAAATATCATGTCT 810  
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40  
DB 811 CAAATTCTGACATGCACACCCAGGGCGTGGCCCAACAAAGATAACTGTGTGCAACGAA 870  
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60  
DB 871 ACGAAGCAAAATACACCAACACGTAGGATAAAACACATGAGATGAAAGCATGCTCACT 930  
QY 61 IleAlaGlyLysSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaPro 80  
DB 931 ATTTGCTGGTCTTGGAAAGAACAAACTTCTGACTACAAAACAGAAACACACTCTCA 990  
QY 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100  
DB 991 CAACAGGAGGTGTGCTGGGAAAGATCCACGAAGTCAAAAGCAGAGTGTGATCT 1050  
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
DB 1051 GGAGTATCAGGAGCGGCTGCTACTACAGCATCAAAATACTGCAACAAAAATTGCTATGCA 1110  
QY 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuInSerLeu 140  
DB 1111 ACCCTATTGAAGAGGCGGCAAAAGTATGAGTCTTACCTTATGATCTCAATCAAGCCCTC 1170  
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSer 160  
DB 1171 AGTGGCGGCAAAATGAAGATGCAAGCGGTGTGTGTCGCGCTCTCAGGAAAGT 1230  
QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
DB 1231 TCGGGTTCGCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290  
QY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyValAlaThr 200  
DB 1291 GAGGTATCGAAATCGGATCGCGCTGTGTAAAGCAATTCAGCATGTGGAGAACCCACA 1350  
QY 201 LysSerAlaLeuSerAsnTyraLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
DB 1351 AATCTGCTTATCTAATCATGACAGTACACAAAGCAAGCAAGCAACAAATTAACATA 1410  
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleGlnMetLys 240  
DB 1411 GGTCTAGAAACCAAGCGATTAATAATTCGATTAAGAAGCAAGAAATATACCAAGATGAAG 1470

QY 331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340  
Db 1991 -----GTGGCGGAGGCGCTCTGGGTCATTTTCGGCAGACCGCTTCCGTGGAGC 2041  
QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
Db 2042 GCGACGATGATGCGGCTGCGGTGCGATTCGGATCTTGACGCCCTCGGTGCAAGCC 2101  
QY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspPhe 380  
Db 2102 TTCGTCACTGTGCGCGC-----ACCAACAGTTTC 2131  
QY 381 ProLysLeuSerLysValIleSerSerLeu----- 390  
Db 2132 GCGGAGACGAGCGCCATTATGCGCGCATGCGCGCGCGCGCTGAGCTTCCTGCT 2191  
QY 391 -----ThrSerLysTrpValThrValGlyValGlyValValAlaAlaProAla 407  
Db 2192 GCGTTCGCGAGCGGAGCGCTGATGCGCTTCGCCATTTGATTTCTTCGCTTCGCGGCGC 2251  
QY 408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPhe 427  
Db 2252 ATCGGGATGCGCGCTGCGAGCGCATGCTCCAGCGAGGTAGATGACGACCATCAGGGA 2311  
QY 428 GlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
Db 2312 CACCTTCA--GGATCGTCGCGGCTTACACAGCCTAAGCTTCGATC 2356  
RESULT 20  
LOCUS E12523 5658 bp DNA linear PAT 27-APR-1998  
DEFINITION Nucleotide sequence of pcPN533 alpha that contains Chlamydia  
pneumoniae antigen gene.  
ACCESSION E12523  
VERSION E12523.1 GI:3251356  
KEYWORDS JP 1997009974-A/4.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 5658)  
AUTHORS Izutsu, H. and Matsumoto, A.  
TITLE ANTISEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING  
FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT  
CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE  
Patent: JP 1997009974-A 4 14-JAN-1997;  
JOURNAL HITACHI CHEM CO LTD  
COMMENT OS None  
OC Artificial sequences.  
PN JP 1997009974-A/4  
PD 14-JAN-1997  
PE 20-SEP-1995 JP 1995242095  
PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI  
IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC C12N15/09, C07H21/04, C07K14/295, C07K16/12, C12N1/21, C12N15/02,  
PC C12P21/08,  
PC C12P21/08,  
PC G01N33/53, G01N33/571//A61K39/118, A61K39/118, A61K39/118, PC  
(C12N1/21, C12R1:19),  
PC (C12P21/08, C12R1:91);  
CC strandedness: Double;  
CC topology: Circular;  
FH Key Location/Qualifiers  
FT source 1. 5658  
FT FT Location/Qualifiers  
FEATURES  
source 1. 5658  
Location/Qualifiers  
1. 5658  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 1446 a 1477 c 1428 g 1307 t  
ORIGIN  
Alignment Scores:

Pred. No.: 1,68e-60 Length: 5658  
Score: 1282.00 Matches: 315  
Percent Similarity: 62.59% Conservative: 33  
Best Local Similarity: 56.65% Mismatches: 75  
Query Match: 54.12% Indels: 135  
DB: 6 Gaps: 10  
US-09-889-314-2 (1-496) x E12523 (1-5658)  
QY 1 AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
Db 751 GATACAAACATGCTATTATTCATCTTCTTCAGACCTACATCAAAAAATATCATCTCT 810  
QY 21 GlnValLeuThrSerTrpProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40  
Db 811 CAGTTTGATCGATCGACACCCAGGGCGTCCCAACAAATTAAGCTGTCTGGCAAGGAA 870  
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60  
Db 871 ACGAAGCAATATACGCAACACGTCAGGTAAACACTAGATGAAACGATGCCACT 930  
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80  
Db 931 ATTGCTGTGCTTCTCGAAAAAGACAAACTTCCCTGACTACAAAAACAGAAACAGTCCA 990  
QY 81 GlnGlnGlnValAlaAlaGlyLysGlnSerSerSerGlnSerGlnLysAlaGlyAlaAspThr 100  
Db 991 CAACAGGAGCTGTGCTGGGAAAAATCCTCAGAAAGTCAAAAGCGAGTGTGATCT 1050  
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaPheGln 120  
Db 1051 GGAGTATCAGAGGCGCTGTACTACAGCATCAATCTGCACAAATAATTGCTTTCGAC 1110  
QY 121 ThrSerIleGluGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeu 140  
Db 1111 ACCCTATTGAAAGGCGACCAAAAGTATGAGTCACTTACGATCACTTCAAGCTTC 1170  
QY 141 SerAlaIleGlnMetLysGluValGlnAlaValValAlaAlaLeuSerGlyLysSer 160  
Db 1171 AGTCCCGCAATGAAAGAGTGAAGCGGTGTGTCTGCTGCTCAAGGAAAAAGT 1230  
QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
Db 1231 TCGGCTCCGCAAAATGGAACACCTGAGCTGCCCAAGCCGGGGTGACACCAAGTCA 1290  
QY 181 GlnValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr 200  
Db 1291 GAGGTATCGAAATCGGACTCGCGCTTCTTAAGCAATTCAGCATTCGGAAGAGCCACA 1350  
QY 201 LysSerAlaLeuSerAsnLysAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
Db 1351 AAATCTGCTTATCTTACTATGCAAGTACACAGCACAGACAGCAACAAATAAATCACTA 1410  
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysArgGlnLysLeu 240  
Db 1411 GGTTAGAAAAAGCAGAGGATTAATAATGATTAAGAAAGAGATTCACCAAGATGAAG 1470  
QY 241 AlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260  
Db 1471 GCTGCCAAGACAGAGTCTTAAGATCTGGAAGGAACAATGATCTGCAATACGTGATG 1530  
QY 261 IleAlaValSerValAlaIleThrVal-----IleSer 271  
Db 1531 ATCGCGAA--GGGTTGCAATTCATGCCATGGGCGCTTATTAATTAACTCAGACATCCAG 1589  
QY 272 IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly 289  
Db 1590 ATCAATACGATGATCCCTTACGCCGAGCATGCTGGCCGCAATCACCAGCGCACAGGT 1649  
QY 290 AlaAlaValGlyAlaAla----- 295  
Db 1650 GCGGTGTCTGGCGCTA--TATCGCGCAGATCACCAGATGGGAAGATCGGGCTCGCACTT 1708

Oy 344 SerGlyIleValPheIleThrLeuValIleAlaIleAlaIleSerLys 363  
 Db 1504 TCTGGAATAAAGCATTTATCAAAACCTTAGCAAGGATTCGCAAGCCATTTCTAAA 1563  
 Oy 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
 Db 1564 GGAATCTCTAAGGTTTTCGCTTAAGGGAAGCTCAATGATTCGCAAGACTTCCCAAGCTC 1623  
 Oy 384 SerLysValIleSerSerLeuThrSerLysTrpValIleThrValGlyValGlyValVal 403  
 Db 1624 TCGAAGATCACTCTCTCTTACCAAGTAATGGCTACAGGTTGGGTTGGGTTTACTT 1683  
 Oy 404 AlaIleProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 423  
 Db 1684 GCGGCGCTGCTCTCGTTAAAGGATTAATGCAATGCAAGCTCTCGAGATGCAACAAAC 1743  
 Oy 424 ValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443  
 Db 1744 GTGCTCAATTTACAGAAAGATCGCAAAACCTGCAAGCTCGGCTGATATGATTTCTATG 1803  
 Oy 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
 Db 1804 TTCACTCAATTTGGCAACAGCAAGTAATGGCTCAAAACAAACAGCGAGTCTAAT 1863  
 Oy 464 GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaIleAlaIle 483  
 Db 1864 GAAATGACTCAAAAGCATRCCAAAGCTGGGCTCAATCTTAAAGCGTATGCCCAATC 1923  
 Oy 484 SerGlyAlaIleAlaGlyAla 490  
 Db 1924 AGCGAGCATCGCTGGCGCA 1944

RESULT 19  
 ARI22110 5658 bp DNA linear PAT 16-MAY-2001  
 LOCUS Sequence 10 from patent US 6165478.  
 DEFINITION ARI22110  
 ACCESSION ARI22110.1 GI:14106427  
 VERSION ARI22110.1 GI:14106427  
 KEYWORDS  
 ORGANISM Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 5658)  
 AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.  
 TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
 JOURNAL Patent: US 6165478-A 10 26-DEC-2000;  
 FEATURES Location/Qualifiers  
 source 1..5658  
 BASE COUNT 1446 a 1477 c 1428 g 1307 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,68e-60 Length: 5658  
 Score: 1282.00 Matches: 315  
 Percent Similarity: 62.59% Conservative: 33  
 Best Local Similarity: 56.65% Mismatches: 75  
 Query Match: 54.12% Indels: 135  
 DB: 6 Gaps: 10

US-09-889-314-2 (1-496) x ARI22110 (1-5658)

Oy 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
 Db 751 GATACAAACATGTATTCATCTTCATCAGACCTGACATCAAAATAATCATGTCT 810  
 Oy 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40  
 Db 811 CAAGTTCATCATGACACCCAGGCGCTGCCCAACAGATTAAGCTGTGCGCAACGAA 870  
 Oy 41 ThrLysGlnIleGlnIleThrArgGlnGlyLysAsnThrGlnMetGlnSerAspIleThr 60

Db 871 ACGAAGCAATACAGCAAAACGTCAGGGTAAAAACACTGATGGAAGGATGCCACT 930  
 Oy 61 IleAlaGlyIleSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80  
 Db 931 ATTCGTGGTCTCTCGGAAAGACAAAACCTTCCTGACTACAAAACGAAACAGCTCA 990  
 Oy 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100  
 Db 991 CAACGAGGAGTTGCTGCGGGAAGAATCCACAGAAATCAAAAGGACGATGATCT 1050  
 Oy 101 GlyValSerGlyAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
 Db 1051 GGAGTATCAGAGCGGCTGCTACTACACATCAATACCTGCAACAAAATGCTATGCG 1110  
 Oy 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeu 140  
 Db 1111 ACCTCTATTGAAGAGCGAGCAAAAGTGAAGTCTACCTTAAGTCACTCAAGGCTC 1170  
 Oy 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaIleAlaLeuSerGlyLysSer 160  
 Db 1171 AGTCCGCGCAATCAAGAAAGTCAACGCGTGTGTTGCTGCCCTCTCAGGGAAGT 1230  
 Oy 161 SerGlySerAlaLysLeuGluThrProGlnLeuProLysProGlyValThrProArgSer 180  
 Db 1231 TCGGGTCCGCAAAATTTGAAAACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290  
 Oy 181 GluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnIleThrLeuGlyGluAlaThr 200  
 Db 1291 GAGTTATCGAAATGCGACTGCGGCTGCTAAAGCAATTCAGACATTTGGAGAGCCCA 1350  
 Oy 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
 Db 1351 AAATCGCTTATCTAATCAATGCAAGTACACAGACACAGCAAGCAACCAATTAACCTA 1410  
 Oy 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGluThrGlnLeuMetLys 240  
 Db 1411 GGTCTAGAAAAGCAAGCATTAATAAATCGATTAAGCAAGAAATACCAAGATCAAG 1470  
 Oy 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260  
 Db 1471 GCTCCGAACAGAACTGTAAGATCTGACAGGACAAATGATGATCAATATCTGATG 1530  
 Oy 261 IleAlaValSerValAlaIleThrVal-----IleSer 271  
 Db 1531 ATCCGCAAG-GGGGTTCGAATTCCTCCATGCGGGGCCCTTAATTAATTAACGAGATCCAG 1589  
 Oy 272 IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly 289  
 Db 1590 ATCTAATGATGATCTCTCTACGCGGACGATCGTGGCGGATCACCGGCGCCACAGGT 1649  
 Oy 290 AlaAlaValGlyAlaAla----- 295  
 Db 1650 GCGGTTCCTGGGCGCTA-TATCGCCGACATCACGATGGGGAAGATCGGCTCGCCACTT 1708  
 Oy 296 -----AlaAlaGlyLysAla 300  
 Db 1709 CCGGCTCATGAGCGCTGTGTTGGCGGTGGATAGTGGCAGGCGCCGTGGCGGGGACTG 1768  
 Oy 301 AlaGlyAla-----AlaAlaAla----- 306  
 Db 1769 TTGGGCGCATCTCTCTGATGACACCATTCCTTGGGGGGGGGTCTCACGGCCTCAAC 1828  
 Oy 306 ----- 306  
 Db 1829 CTACTACTGGGCTGCTCTCTAATGAGAGTGCATTAAGGAGAGCGTCGACCGATGCC 1888  
 Oy 307 -----ThrValAla 310  
 Db 1889 TTGACAGCCTTCAACCCAGTCACTCTTCGCGTGGCGCGGGGATGATCACTATCTGCGC 1948  
 Oy 311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330  
 Db 1949 GCACCTATGACGCTCTCTTATTATGCAACTGCTGAGACAG----- 1990



|||||  
Db 1684 GCGGCGCTGCTCGTAAGGATATGCAATGACGCTCGAGATGCAACAAAC 1743  
QY 424 ValAlaGlnheGlnySgluValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443  
Db 1744 GTGCTCAATTGAGAAAGATGCGAAACCTGAGCTCGCTGATATGATTTCTATG 1803  
QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnPheTrpGlnSerAsn 463  
Db 1804 TTGACTCAATTTTGGCAGACGCAAGTAAATTCCTCAAAACAAACAGCGAGCTTAAT 1863  
QY 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaIle 483  
Db 1864 GAATGACTCAAAAGCTACCAAGCTGGCGCTCAAAATCTTAAGCGGTATGCCCAATC 1923  
QY 484 SerGlyAlaIleAlaGlyAla 490  
Db 1924 AGCGAGCCATGCTGCGCA 1944  
RESULT 18  
E12538  
LOCUS E12538 1947 bp DNA linear PAT 27-APR-1998  
DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of  
Chlamydia pneumoniae.  
ACCESSION E12538  
VERSION E12538.1 GI:3251371  
KEYWORDS JP 1997015243-A/1.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1947)  
AUTHORS Izutsu, H., Obata, K. and Matsumoto, A.  
TITLE MEASURING METHOD FOR ANTI-CHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT  
THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA  
INFECTION  
JOURNAL Patent: JP 1997015243-A 1 17-JAN-1997;  
HITACHI CHEM CO LTD  
COMMENT OS None  
OC Artificial sequences.  
PN JP 1997015243-A/1  
PD 17-JAN-1997  
PF 15-MAR-1996 JP 1996058609  
PI 28-APR-1995 JP 95P 106009  
PR IZUTSU HIROSHI, OBATA KAZUHIKO, MATSUMOTO AKIRA PC  
G01N33/569, C12N1/21, C12N15/02, C12P21/02, C12P21/08, PC  
G01N33/53  
PC G01N33/531, G01N33/571//A61K49/00, C07K14/245, C07K14/39, PC  
C07K16/12, C07K19/00,  
PC (C12N1/21, C12R1:19), (C12N15/09, C12R1:01), (C12P21/02, C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key Location/Qualifiers  
FT source 1. 1947  
FT /organism='Artificial sequences' FT  
FT mat\_peptide 1. 1947  
FT /product='fusion protein of DHFR and antigen  
FT peptide'  
FT misc\_feature 1. 480  
FT /note='DHFR coding region'  
FT 484..1947  
FT misc\_feature /note='antigen peptide coding region'.  
FEATURES  
source Location/Qualifiers  
1. 1947  
/organism='unidentified'  
/db\_xref='taxon:32644'  
BASE COUNT 591 a 449 c 486 g 421 t  
ORIGIN  
Alignment Scores: 5.71e-117 Length: 1947  
Pred. No.: 2316.00 Matches: 487  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00%

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
DB: 6 Gaps: 0  
US-09-889-314-2 (1-496) x E12538 (1-1947)  
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
Db 484 ATGCTATTTCATCTTCTTACGACCTGACCAATCAAAAAATATCTCATGTCTCAAGTTCTG 543  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
Db 544 ACATCGACACCCAGCGGCGTCCCAAGATTAAGCTGCTGCGCAACGAAGCAAGCA 603  
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63  
Db 604 ATACAGCAAAACAGCTCAGGCTGTAACACACAGATGTAAGCAATCCACTATGCTGCTG 663  
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
Db 664 GCTTCTGGAAGAACCAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723  
QY 84 ValAlaAlaGlyLysGluSerSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
Db 724 GTTGCTGCTGGAAGAAATCTCAGAAAGTCAAAAGCAGCAGCTGCTGATCTGAGATATCA 783  
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
Db 784 GGAGCGGCTGCTACTACAGATCAAAATACGCAACAAAAATTCCTATGCGACCTCTATT 843  
QY 124 GlnGlnLysSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAla 143  
Db 844 GAAGAGCGGCAAGCAAGTACGAGCTCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 903  
QY 144 GluMetLysGluValGluAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163  
Db 904 CAATGAAAGAAAGTCAAGGCGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963  
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183  
Db 964 GCAAAATGGAACACCTGAGCTCCCAACCCGGGGTGACCAAGATCAGAGGTATTC 1023  
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyLysAlaThrLysSerAla 203  
Db 1024 GAATGCGAGTCCGCGCTGCTGTAAGCATTCACACATTTGGGGAACCCCAAAATCTGCC 1083  
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223  
Db 1084 TTATCTTAACATATGCAAGTACACAGCAACAGCACCAACCAAAATTAAGTCTTAAGAA 1143  
QY 224 LysGlnAlaIleLysIleAspLysGluArgGlnGlnIleGlnIleMetLysAlaIleGlu 243  
Db 1144 AAGCAAGCGATTAATAATCGTAAGAAAGAGAAATACCAAGATGAAGCGTGCAGAA 1203  
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 1204 CAAAGCTTAATATCTCGAAGAACATGATACGTCAATCTGTGATGATGCGGGTT 1263  
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
Db 1264 TCTGTGCTCATTTACAGTATTCTATGTTGCTGATTTTTCATAGCGGAGCTGAGATC 1323  
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaIleGlyAlaAlaGlyAla 303  
Db 1324 GCTGAGCTGCTGCGGAGCTGCTGAGTGCAGCGGCGAGCTGAGAGTCAAGAGAGCT 1383  
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323  
Db 1384 GCTGCCGACCAACGAGTACGACACAAATTTACGTTCAAGCTGTTTCCAAAGCGGTGA 1443  
QY 324 GlnAlaValIleThrAlaValArgAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
Db 1444 CAAGCTGTTATCACAGCTGTCAGACAAAGCGATACCGCGGCTTAAAGCGGCTGTCA 1503



DB 1924 AGCGAGCCATGCTGGCGCA 1944

RESULT 17

LOCUS E12525

DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.

ACCESSION E12525

VERSION E12525.1 GI:3251358

KEYWORDS JP 1997009976-A/L.

SOURCE unclassified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1947)

AUTHORS Izutsu, H., Obara, K. and Matsumoto, A.

TITLE DIHYDROFOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE ANTIBODY

JOURNAL Patent: JP 1997009976-A 1 14-JAN-1997;

HITACHI CHEM CO LTD

COMMENT OS None

OC Artificial sequences.

PN JP 1997009976-A/1

PD 14-JAN-1997

PF 14-MAR-1996 JP 1996057409

PR 28-APR-1995 JP 95P 106006

PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC

C12N15/09, C07H21/04, C07K14/295, C07K14/47, C07K19/00, C12N1/21, PC

C12P21/02,

PC GOIN33/569, GOIN33/571//A61K39/395, A61K49/00, C12P21/08, PC

(C12N15/09, C12R1:01),

PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91),

PC (C12P21/08,

PC C12R1:91);

CC Strandedness: Double;

CC Topology: Linear;

FH Key Location/Qualifiers

FT source 1. 1947

FT mat\_peptide 1. 1947

FT /product='fusion protein of DHFR and antigen peptide'

FT misc\_feature 1. 480

FT /note='DHFR coding region'

FT misc\_feature 484..1947

FT /note='antigen peptide coding region'.

FEATURES

source 1..1947

Location/Qualifiers

/organism='unidentified'

/db\_xref='taxon:32644'

BASE COUNT 591 a 449 c 486 g 421 t

ORIGIN

Alignment Scores:

Pred. No.: 5,71e-117 Length: 1947

Score: 2316.00 Matches: 487

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 97.76% Indels: 0

DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12525 (1-1947)

OY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnIysAsnIleMetSerGlnValIleu 23

DB 484 ATGCTGATTTCACTCTCTCAGAGACTGACATCAAAAAATATCATGTCACAGTTCTG 543

OY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyIysAsnGluThrIysGln 43

DB 544 ACATGCACACCCAGGCGCTGCCCAACAAGATTAACGTCTGTGGCAACGAAGCAAGCA 603

OY 44 ILeGlnGlnThrArgGlnGlnIlysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63

DB 604 ATACAGCAACACCGTCAGGCTAAACACTGAGATGAGAAAGCGATGCCATTTGCTGCT 663

OY 64 AlaSerGlyLysAspLysThrSerSerThrThrIysThrGluThrAlaProGlnGlnGly 83

DB 664 GCTTCTGAAAAGACAAAACCTTCCTGACTACAAAACGAAACAGACGCTCCACACAGGGA 723

OY 84 ValAlaIleAlaGlyLysGlnSerSerGlnIysAlaGlyAlaAspThrGlyValSer 103

DB 724 GTTGCTGTGGGAAAGAACTCCACAAAGTCAAAAGGCGGTGTGATCTGAGTATCA 783

OY 104 GlyAlaAlaIleAlaThrThrAlaSerAsnThrAlaThrIysIleAlaMetGlnThrSerIle 123

DB 784 GGAGCGGCTGCTACTACACATCAATTAATCTGCACAAAATTTGCTATGACACCTTATT 843

OY 124 GluGluAlaSerLysSerMetGlnSerThrIleGlnGlnSerLeuGlnSerLeuSerAla 143

DB 844 GAAGAGCGCAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCCTCAGTGCGCG 903

OY 144 GlnMetLysGluValGluAlaValValAlaIleAlaLeuSerGlyLysSerSerGlySer 163

DB 904 CAAATGAAAAGAGTCAACCGGCTGTGTGCTGCCCTCCAGGGAAGTTTCGGGTTCC 963

OY 164 AlaLysLeuGluThrProGlnLeuProLysProGlnValThrProArgSerGluValIle 183

DB 964 GCAAAATTGGAAACACCTGAGTCCCCAGGCGGGGTACACCAAGATCAAGAGTTATC 1023

OY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203

DB 1024 GAATCGAGCATCGCGCTGCTTAAGCAATTCAGACTTGGAGAAAGCCAAATCTGCC 1083

OY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGlu 223

DB 1084 TTATCTTAATCTATGACAGTACACAGCACAGGACCAAACTAATAGTCTAGAA 1143

OY 224 LysGlnAlaIleLeuIleAspLysGluArgGlnGluGluThrGlnMetLysAlaAlaGlu 243

DB 1144 AGCAAGCGATTAATAATCATTAAGACAGAGAAAGATACCAAGATGAAAGGCTGCCGA 1203

OY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263

DB 1204 CAGAAGCTTAAGATCTCAAGAGCAACAAATGATACGTCAATAGCTGATGATCGCGGTT 1263

OY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283

DB 1264 TCTGTTCATTAACAGTATTTCTATTGTTGCTGCTATTTTACATGCGGACTGACTC 1323

OY 284 AlaGlyLeuAlaAlaGlyValAlaValGlyValAlaAlaGlyValAlaAlaGlyVal 303

DB 1324 GCTGGCTGCTGCGGAGCTGCTGAGTGCAGCGGCGAGCTGAGGAGTCCACAGAGGCT 1383

OY 304 AlaAlaAlaIleThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323

DB 1384 GCTGCGGCAACACGCTACCAACAAATTAACAGTTCAAGCTGTGTGCCAACGCGGAA 1443

OY 324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaValLys 343

DB 1444 CAAGCTGTATACAGCTGCTCAGCAAGCAGACCGCGCTATTAACCGCTGTCAAA 1503

OY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363

DB 1504 TCTGGATTAAGCATTTTATCAAAACCTTAGTCAAAAGCATTTGCCAAAGCCATTTCTAAA 1563

OY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383

DB 1564 GGAATCTTAAGGTTTTCCTTAAGGGAACCTCAATGATTCGCAAGAACTTCCCAAGCTC 1623

OY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403

DB 1624 TCGAAAGTATCTGCTCTTACAGTAATAGGTGACAGGTTGGGGTGTGTGAGTT 1683

OY 404 AlaAlaProAlaLeuGlnGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnIleAsn 423

Db 1201 GCGGCGCTCTCGGTAAAGGATTATGCAATTCAGCTCTCGAGATGCACAACAAC 1260  
QY 424 ValAlaInpHeGInLysGluValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443  
Db 1261 GTGCTCAATTTGAGAAAGATCGAACAACCTGAGGCTGCGCTGATATGATTTCTATG 1320  
QY 444 PheThrGlnPheThrProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
Db 1321 TTGACTCAATTTTGGCAACAGGCAAGTAAATTTGCTCAAAACAAACAGGAGCTTAAT 1380  
QY 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIle 483  
Db 1381 GAAATGACTCAAAAAGCTACCAACGCTGCGCTCAAAATCCTTAAGGTATGCGCAATC 1440  
QY 484 SerGlyAlaIleAlaGlyAla 490  
Db 1441 AGCGAGCCATCGCTGCGCA 1461  
RESULT 16  
AR122114  
LOCUS AR122114 1947 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 17 from patent US 6165478.  
ACCESSION AR122114  
VERSION AR122114.1 GI:14106431  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1947)  
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.  
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
JOURNAL Patent: us 6165478-A 17 26-DEC-2000;  
FEATURES  
source location/Qualifiers  
1..1947  
BASE COUNT 591 a 449 c 486 g 421 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.71e-117 length: 1947  
Score: 2316.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
DB: 6 Gaps: 0  
US-09-889-314-2 (1-496) x AR122114 (1-1947)  
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
Db 484 ATGCTATTTCATCTCTTCGAGACGTGACAAATATATCATGTCTCAAGTTCTG 543  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
Db 544 ACATCGACACCCGAGGCGGTGCCCAACAAATAGCTGTGGCAACGAACGAACGA 603  
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
Db 604 ATTAACGAACACGTCAGGATTAACACTGAGATGCAAAACGATGCCACATTTGCTGAT 663  
QY 64 AlaSerGlyLysAspLysThrSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
Db 664 GCTTCTGGAAGACAAACTCTCTCGACTACAAAACAGAAACAGACGTCACAAACAGGA 723  
QY 84 ValAlaAlaGlyLysGluSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
Db 724 GTTCTGCTGCGAAGAACTCTCAAGAAAGTCAAAAGCAGCTGCTGATCTGAGTATCA 783  
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
Db 784 GGAGCGGCTGCTACTACGACATCAAAATATCTGCACAAAATTTGCTATGACAGCTTAT 843  
QY 124 GluGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerAlaAla 143

Db 844 GAAGAGCGAGCAAAAGTATGAGTACTACCTTAGAGTCACATTCAAAAGCTCAGTGCAGCG 903  
QY 144 GlnMetLysGluValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer 163  
Db 904 CAAATGAAAGAAAGTCGAAGGGGTGTGTGTCGCCCTCTCAGGAAAGTATCGGGGTTC 963  
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183  
Db 964 GCAAAATTTGCAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGTTATC 1023  
QY 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyLysAlaThrLysSerAla 203  
Db 1024 GAAATCGAGCTCGCGCTTGCTTAAGCAATTCAGACATTTGGAGAGAACCAAAAATCTGCC 1083  
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223  
Db 1084 TTATCTAATATGCAATGACACAGCAAGCAGACACCAACAAATTAACATGATGTTAGAA 1143  
QY 224 LysGlnAlaIleLysIleAspLysGluArgGlnGlnThrGlnGlnMetLysAlaAlaGlu 243  
Db 1144 AAGCAGCGATTAATATCGATTAAGAAACGAGAGATATCCAAAGATCAAGGCTGCCGAA 1203  
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 1204 CAGAAGTCTTAAGATCTCGAAGGAACAATGATGATCTGCAATACTGTGATGATCGCGTT 1263  
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
Db 1264 TCTGTTCGCAATTAAGTATTTCTATTGTGCTCTCTTTTAAATCGGAGCGTGCAGCTC 1323  
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303  
Db 1324 GCTGGACTCGCTCGCGGAGCTGCTGTAGTGTCACGCGGAGCTGAGGTGCACGAGAGCT 1383  
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValValGlnAlaValLys 323  
Db 1384 GCTCGCGCAACCAAGCTGACGACCAACAATATACATTCAGCTGTTGTCACAGCGGTGAA 1443  
QY 324 GlnAlaValAlaIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
Db 1444 CAACGCTGTAAACAGCTGTGACGACCAAGCGATCCGCGGTATTAAGCGCGCTGATAA 1503  
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
Db 1504 TCTGGAATTAAGCATTTATTAACAACTTACTCAACGAGATTCGCAAAAGCATTTCTAAA 1563  
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
Db 1564 GGAATCTCTAAGTTTTCGCTAAAGGAACCTCAAAATGATTCGCAAGAACTTCCCAAGCTC 1623  
QY 384 SerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValValVal 403  
Db 1624 TCGAAAGTCATCTGCTCTTAACCAATTAATGGGTCAAGGGTGGGGTGGGTGTGTAGTT 1683  
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423  
Db 1684 GCGGCGCTGCTCGGTAAAGGATTAATGCAAAATGCAAGCTCTCGGAGATGCACAACAAC 1743  
QY 424 ValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
Db 1744 GTCCCTCAATTTCAAGAAAGATGCGAAMAACCTGCAAGCTGCGCTGATATATTTCTATG 1803  
QY 444 PheThrGlnPheThrProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
Db 1804 TTCACTCAATTTTGGCAACAGCAAGTAAATTCCTCAAAAACAAACAGGAGCTTAAT 1863  
QY 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIle 483  
Db 1864 GAAATGACTCAAAAAGCTACCAACGCTGCGCTCAAAATCCTTAAGGTATGCGCGCAATC 1923  
QY 484 SerGlyAlaIleAlaGlyAla 490  
|||||

Db 1321 TTCACGATTTTGGCAACAGGCAAGTAAATTCCTCAAAACAAACAGCGACTTAA 1380  
QY 464 GUmethThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaIle 483  
Db 1381 GAAATGACGCAAAAGCTACCAAGCTGGCGCTCAATCTTAAAGCTATGCGCAATC 1440  
QY 484 SerGlyAlaIleAlaGlyAla 490  
Db 1441 AGCGAGCATCTCGTGGCCA 1461  
RESULT 15  
E16635 1464 bp DNA linear PAR 28-JUL-1999  
LOCUS Chlamydia pneumoniae gene for 53kDa-antigen peptide.  
DEFINITION E16635  
ACCESSION E16635  
VERSION E16635.1 GI:5711318  
KEYWORDS JP 198210978-A/1.  
SOURCE JP 198210978-A/1.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1464)  
AUTHORS Izutsu,H. and Hagihara,T.  
TITLE RECOMBINANT VECTOR AND TRANSFORMANT CONTAINING THE SAME, AND  
RECOMBINANT VACUOLAR VIRUS AND ITS PRODUCTION, AND PRODUCTION OF  
CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE  
PATENT: JP 198210978-A 1 11-AUG-1998.  
JOURNAL HITACHI CHEM CO LTD  
COMMENT OS Chlamydia pneumoniae  
PN JP 198210978-A/1  
PD 11-AUG-1998  
PF 31-JAN-1997 JP 1997018523  
PI IZUTSU HITROSHI, HAGIMARA TOSHIKATSU  
PC C12N15/09,C07H21/04,C07K14/235,C12N1/21,C12N5/10,C12N7/00, PC  
C12P21/02//  
PC A61K39/00,A61K39/118,C12Q1/68,(C12N15/09,C12R1:01),(C12N1/21,  
PC C12R1:19),  
PC (C12N5/10,C12R1:91),(C12N7/00,C12R1:92),(C12P21/02,C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1.1464  
FT /strain='YK-41' /organism='Chlamydia pneumoniae' FT  
FT mat\_peptide 1.1464 /tissue\_type='elementary bodies' FT  
FT /db\_xref='taxon:32644' /product='53kDa-antigen peptide'.  
FEATURES  
source Location/Qualifiers  
1.1464  
/db\_xref='taxon:32644'  
BASE COUNT 470 a 332 c 348 g 314 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.05e-117 Length: 1464  
Score: 2316.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
Gaps: 0  
US-09-889-314-2 (1-496) x E16635 (1-1464)  
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
Db 1 AAGCTATATTCATCTCTTCAGAGCACTGACAAATCAAAAAATATCATGCTCAAGTTCTG 60  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerSerGlyAsnGlnThrLysGln 43  
Db 61 ACATTCAGACCCCGAGGCGGCCCAAGATTAAGCTGCTGCGCAACGAAACGAAAGCA 120  
QY 44 IleGlnIleThrArgGlnGlnLysAsnThrGlnMetGlnUserAspIleThrIleAlaGly 63

Db 121 ATACAGCAAAACGCTCAGGCTAAAAACACTGAGATGGAAGGATGCCACTTGTGCTGT 180  
QY 64 AlaserGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
Db 181 GCTCTGGAAAAAGACAAAACCTTCCTCGACTACAAAACGAAAACAGCTCCACAAAGGAA 240  
QY 84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyAlaSer 103  
Db 241 GTTGCTGCTGGGAAAGATCTTCAGAAAGTCAAAAGCGAGGTGATCTGAGATATCA 300  
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
Db 301 GGACGGCGCTGCTACTACACATCAAAATAGTGCACAAAAATTCATGACAGACTCTAT 360  
QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuAlaAla 143  
Db 361 GAGAGCGCAGCAAAAGTATGAGTCTACCTTACGATCAAAAGCTCAGTGCCTGCGG 420  
QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163  
Db 421 CAAATGAAGAAGTCGAACGGGTGTTGTTGCTGCTCAGGGAAGTTGCGGCTCC 480  
QY 164 AlalysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183  
Db 481 GCAAAATTTGGAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGTTATC 540  
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203  
Db 541 GAAATCGGACCTCGCGCTCTTAAGCAATTCAGACATTGGAGAAAGCCCAAAATCTGCC 600  
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAspLysLeuGlnGlu 223  
Db 601 TTATCTTAATGACAGTACACAGCAAGCAGCAAGCAAAATTAATCTAGTCTAGAA 660  
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 243  
Db 661 AAGCAAGCATTAATAATCATTAAGAAAGCAAGAAATATCAAGATGAAAGCTGCCGAA 720  
QY 244 GlnLysSerLysAspLeuGlnGlnThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 721 CAGAACTCTTAAGATCTCCAGAGCAAGCAAGATGATGATGATGATGATGATGATGAT 780  
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
Db 781 TCTGTTGCCATTAACATTAATTTCTATTGTTGCTGCTATTATTTTACATGCGAGCTGACTC 840  
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAla 303  
Db 841 GCGGACTCGCGCGGAGCTGCTGATGAGTGCAGCGGAGCTGAGAGTGCAGAGAGCT 900  
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323  
Db 901 GCTGCGCAACACCGCTACCAACCAAAATTTACAGTCTGTTCTCCAAAGCGGGGAAA 960  
QY 324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
Db 961 CAAAGCTGTTATCACACCTGTCAGCAAGGATACCGCGGCTATTAAGCGGCTGCAAA 1020  
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
Db 1021 TCTGGAATAAAGCATTTATCAAAACTTTAGTCAAAAGCGATTCGCAAAACCAATTTCTA 1080  
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
Db 1081 GGAATCTTAAGGTTTTCCTTAAGGAACTCAAAATGATTCGGAAGAACTTCCCAAGCTC 1140  
QY 384 SerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValVal 403  
Db 1141 TCGAAAGTATCTCGTCTTACCACTAAATGGGTACGGTTGGGTTGTAGTT 1200  
QY 404 AlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423

QY 484 SerGlyAlaIleAlaGlyAla 490  
Db 1441 AGCGAGCCATCGCTGGCGCA 1461  
RESULT 14  
E12541  
LOCUS E12541 1464 bp DNA linear PAT 27-APR-1998  
DEFINITION DNA encoding an antigen peptide Chlamydia pneumoniae.  
ACCESSION E12541  
VERSION E12541.1 GI:3251374  
KEYWORDS JP 1997015244-A/1.  
SOURCE unidentifed.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1464)  
AUTHORS Iizusu, H. and Matsumoto, A.  
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION  
JOURNAL Patent: JP 1997015244-A 1 17-JAN-1997;  
HITACHI CHEM CO LTD  
COMMENT OS Chlamydia pneumoniae  
PN JP 1997015244-A/1  
PD 17-JAN-1997  
PF 14-MAR-1996 JP 1996057410  
PR 28-APR-1995 JP 95P 106011  
PI IZUTSU HIROSHI, MATSUMOTO AKIRA  
PC GOIN33/571, COTK14/295, COTK16/12, GOIN33/53,  
GOIN33/569//A61K49/00, C12N15/09;  
CC strandedness: Double;  
CC topology: linear;  
FH Key  
FT source 1. 1464  
FT Location/Qualifiers  
FT /strain='YK41',/organism='Chlamydia pneumoniae' FT  
FT mat\_peptide 1. 1464  
FT /product='antigen peptide',  
FT Location/Qualifiers  
BASE COUNT 470 a 331 c 348 g 315 t  
ORIGIN  
FEATURES  
source 1. 1464  
/organism='unidentifed'  
/db\_xref='taxon:32644'  
Alignment Scores:  
Pred. No.: 4.05e-117 Length: 1464  
Score: 2316.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
Gaps: 0  
US-09-889-314-2 (1-496) x E12541 (1-1464)  
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
Db 1 ATGTCTATTTCATCTTCTTACGAGCTGACATCAATCAAAAAATATCATGTCTCAAGTTCTG 60  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
Db 61 ACATGCACACCCGAGCGCGGCCCAACAAGATAGCTGTCTGGCAACGAAGCAAGCA 120  
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnUserAspAlaThrIleAlaGly 63  
Db 121 ATTCACCAACACGTCAGGTAAACAACTGACATGACATGAAGCAAGCATCCACTATTCTGCTG 180  
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGln 83  
Db 181 GCTTCTGGAAAAACAAACTTCTCTGATACCAAAAAACGAACACCTCCACACAGGGA 240  
QY 84 ValAlaIleGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103

Db 241 GTTCTGCTGGGAAGATCTTCAGAAAGTCAGAAAGCAGGTGCTGATCTGAGATTCACA 300  
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
Db 301 GGAGCGCTGCTACTACAGCATCAATATGTCACAAATAATGCTATGACAGACTCTATT 360  
QY 124 GlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
Db 361 GAAGAGCGAGCAAAATATGAGTACTTACCTTACGTCACCTTCAAAAGCTCAGCGCGC 420  
QY 144 GlnMetLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163  
Db 421 CAATGAAAGACGCGAAGCGGCTGTTGCTGCTGCTCTCAGAGAAAAGTTCGGGTTC 480  
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlyValIle 183  
Db 481 GCAAAATTGGAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGTTATC 540  
QY 184 GlnIleGlyLeuAlaLeuAlaIleAlaIleGlnThrLeuGlnGlnAlaThrLysSerAla 203  
Db 541 GAAATCGAGCTCGGCTTGTCTAAAGCAATTCAGACATTCGGAAGCCACAAATCTGCC 600  
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGln 223  
Db 601 TTATCTACTATGCAATGATACACAGCAAGCAGACCAACCAATTAAGTCTTGAGA 660  
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnMetLysAlaAlaGln 243  
Db 661 AAGCAACGATTAATAATTCGATTAAGAAAGAGAAATACCAAGATGAAGCTGCGCAA 720  
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 721 CAGAAGTCTAAGATCTCGAAGAACATGATGATCTCAATCTGTGATGATCGCGGTT 780  
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
Db 781 TCTGTTGCCATTACAGTATTCTATTGCTGCTGCTATTCTTTCATACCGGAGCTGAGCTC 840  
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303  
Db 841 GCTGAGCTCGCTCGGAGCTGCTGAGTGCAGCGGCGAGCTGAGAGTGCAGAGAGCT 900  
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323  
Db 901 GCTGCCGCAACACGAGTAGCAACAAATTAAGTTCAGCTGTTCCAAAGCGGTGAAA 960  
QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
Db 961 CAGCGCTTATACACAGCTGTCAGACAGCAGATCACCGCGCTTTAAACCGCTGTCAAA 1020  
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
Db 1021 TCTGGAATTAAGCATTTATCAAAACTTATGTCAAAGCATTCGCCAAAGCAATTTCTTAA 1080  
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
Db 1081 GGATCTCTAAGGTTTCCCTAAGGAACTCAAAATGATTCGAAGAATCTCCCAAGCTC 1140  
QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValValVal 403  
Db 1141 TCGAAAGTATCTCGCTCTCTTACCACTAAATGGGTACCGCTTGGGGTTGGGTTGATT 1200  
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423  
Db 1201 GCGGCGCTGCTCTCGTAAAGGATTAATGCAAAATGACAGCTCTCGAGATGCAACAAAAC 1260  
QY 424 ValAlaGlnPheGlnLysLeuValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443  
Db 1261 GTGCTCAATTTTCAGAAAGATCGGAAATCGAGGCTCGGATATGATTTCTATG 1320  
QY 444 PheThrGlnPheThrProlGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463

LOCUS E12535 1464 bp DNA linear PAT 27-APR-1998  
 DEFINITION Chlamydia pneumoniae 53kD-antigen gene.  
 ACCESSION E12535  
 VERSION E12535.1 GI:3251368  
 KEYWORDS JP 1997009999-A/1.  
 SOURCE unclassified.  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 1464)  
 AUTHORS Izutsu, H. and Matsumoto, A.  
 TITLE PROBE AND PRIMER FOR MEASURING CHLAMYDIA PNEUMONIAE GENE.  
 JOURNAL MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE USING THE SAME PROBE OR  
 PRIMER AND REAGENT FOR MEASURING CHLAMYDIA PNEUMONIAE GENE  
 Patent: JP 1997009999-A 1 14-JAN-1997;  
 HITACHI CHEM CO LTD  
 COMMENT OS Chlamydia pneumoniae  
 PN JP 1997009999-A/1  
 PD 14-JAN-1997  
 PE 15-MAR-1996 JP 1996058608  
 PR 28-APR-1995 JP 95P 106008  
 PI IZUTSU HIROSHI, MATSUMOTO AKIRA  
 PC C1201/68, C07H21/04, C12N1/21, C12N15/02, C12N15/09, C12P21/08, PC  
 G01N33/366,  
 PC G01N33/569, G01N33/571//A61K49/00, C07K14/295, C1201/04,  
 (C12N1/21, PC C12R1:19);  
 CC strandedness: Double;  
 CC topology: Linear;  
 FH key Location/Qualifiers  
 FT source 1. 1464  
 FT /strain='YK-41',  
 FT mat\_peptide /product='53kD-antigen peptide',  
 FT 1. 1464  
 FT Location/Qualifiers  
 1. 1464  
 /organism='unclassified'  
 /db\_xref='taxon:32644'  
 BASE COUNT 470 a 331 c 348 g 315 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.05e-117 Length: 1464  
 Score: 2316.00 Matches: 487  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.76% Indels: 0  
 DB: Gaps: 0  
 US-09-889-314-2 (1-496) x E12535 (1-1464)  
 QY 4 MetSerLleSerSerSerSerLyProAspAsnGlnLysAsnLleMetSerGlnValIeu 23  
 DB 1 AAGTCTATTTCATCTCTTCAGAGACCTGCACATCAAAAAATATCATGTCTCAAGTTCTG 60  
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 DB 61 ACATGCACACCCAGGCGCTGCCCAACAAGTAAAGCTCTCTGGCACAAGCAAGCA 120  
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
 DB 121 ATACAGCAAAACGTCAGGCTAAACACTGAGATGGAAGGATGCCACTTTGTGTGT 180  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
 DB 181 GCTTCTGAAAAGACAACTTCCTGACTACAAAAAGCAAGCTCCACACAGGGA 240  
 QY 84 ValAlaIaIaGlnLysGlnLysSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
 DB 241 GTTGGCTGGGAAAGAAATCCTCAGAAAGTCAAAAGCGAGGTCTGATCTGAGATATCA 300  
 QY 104 GlyAlaAlaIaIaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 |||||||

DB 301 GGAGCGGCTGCTACTACAGCATCAAAATACGCCAACAAAAATGCTATGCAGACCTCTATT 360  
 QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuAla 143  
 DB 361 GAAGAGCGGAGCAAAAGTAGTGATCTACCTTAGATCACTTCAAGGCTCAGTCCGG 420  
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaIaIaLeuSerGlyLysSerSerLys 163  
 DB 421 CAAATGAAAGAGTCGAGGCGTTTGTGTGCTGCCCTCTCAGGAAAAAGTTCCGGTTCC 480  
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183  
 DB 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTACACCAAGATCAAGGTTTATC 540  
 QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnValAlaThrLysSerAla 203  
 DB 541 GAAATCGGAGCTGCGCTTGTCTTAAGCAATTCAGACATTTGGGAGAAAGCCCAAAATCTGCC 600  
 QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLys 223  
 DB 601 TTATCTACTATGCAAGTACACAGCAGCAGCAGCAACCAAAATTAATCTAGCTAGAA 660  
 QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysAlaIaIaGln 243  
 DB 661 AAGCAAGCGATAAATGATTAAGAACGAGAAATATCAAGATGAAAGCTTCCGCA 720  
 QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
 DB 721 CAGAAGTCTAAAGATCTGCAAGGAACAAAGATGCTCAATACGTATGATGTCGCGTT 780  
 QY 264 SerValAlaIleThrValIleSerIleValAlaIaIaIlePheThrCysGlyAlaGlyLeu 283  
 DB 781 TCTGTGTCATATACAGTATATTTCTATGTGCTGCTATTTTACATGCGAGCTGAGCTC 840  
 QY 284 AlaGlyLeuAlaIaGlyAlaAlaValGlyAlaIaIaIaIaIaGlyGlyAlaIaGlyAla 303  
 DB 841 GCTGACCTGCTGCGGAGCTCTAGCTGCGCGGCGAGCTGAGGTGCACAGAGCT 900  
 QY 304 AlaAlaIaIaThrThrValAlaThrGlnIleThrValGlnAlaValAlaValLys 323  
 DB 901 GCTGCGGCAACCGGTACCAACACAAATTCAGTTCAGTCTGTGTCCAAAGCGGGA 960  
 QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaIaIaIleLysAlaIaValLys 343  
 DB 961 CAAGCTGTATATACAGCTGTACAGACAGATCACCAGCGCTATTAAGCGGCTGTCAAA 1020  
 QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
 DB 1021 TCTGAAATFAAAGCATTTTATCAAAACTTTAGTCAAAAGCGATTGCCAAAGCCATTCTAAA 1080  
 QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
 DB 1081 GGAATCTCTAAGGTTTTCCTAAGGAACTCAAAATGGAAGTTCAGAAAGCTCCCAAGCTC 1140  
 QY 384 SerLysValIleSerSerLeuThrSerLysTrpValIleThrValGlyValAlaVal 403  
 DB 1141 TCGAAAGTATCTCGTCTCTTACAGTAATGGTCCAGCGTTGGGTTGGGTTGTAGTT 1200  
 QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnLysGlnGln 423  
 DB 1201 GCGGCGCTGCTCTCGGTAAAGGATTATGCAAAATGCAAGCTCTCGAGATGCAACAAAAC 1260  
 QY 424 ValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaIaIaAspMetIleSerMet 443  
 DB 1261 GTTCGTCATTTTCAGAAAGATCGGAAACATGCAAGGCTGCGGTGATATGATTTCTAAG 1320  
 QY 444 PheThrGlnPheThrProGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
 DB 1321 TTCACCTCATTTTGGCAACAGCAAGTAAATTTGCTCTCAAAACCAACGCGAGCTAAT 1380  
 QY 464 GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaIaIaIe 483  
 DB 1381 GAAATGACTCAAAAGCTCCAAAGCTGCGCTCAAAATCTTAAAGCGATATCCCGCAATC 1440  
 |||||||

REFERENCE 1 (bases 1 to 1464)  
 AUTHORS Izutsu, H. and Matsumoto, A.  
 TITLE ANTIGEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE Patent: JP 1997009974-A 1 14-JAN-1997;  
 JOURNAL HIRACHI CHEM CO LTD  
 COMMENT  
 OS Chlamydia pneumoniae  
 PN JP 1997009974-A/1  
 PD 14-JAN-1997  
 PE 20-SEP-1995 JP 1995242095  
 PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI  
 IZUTSU HIROSHI, MATSUMOTO AKIRA  
 PC C12N15/09, C07H21/04, C07K14/295, C07K16/12, C12N1/21, C12N15/02, C12P21/08,  
 PC G01N33/53, G01N33/571//A61K39/118, A61K39/118, A61K39/118, PC (C12N1/21, C12R1:19),  
 PC (C12P21/08, C12R1:91);  
 CC strandedness: Double;  
 CC topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1. 1464  
 FT /strain='YK41', /organism='Chlamydia pneumoniae' FT  
 FT mat\_peptide 1. 1464  
 FT /product='antigen peptide'.  
 FEATURES  
 source Location/Qualifiers  
 1. 1464  
 /organism='unidentified'  
 /db\_xref='taxon:32644'  
 BASE COUNT 470 a 331 c 348 g 315 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.05e-117 Length: 1464  
 Score: 2316.00 Matches: 487  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.76% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-889-314-2 (1-496) x E12520 (1-1464)  
 QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
 DB 1 ATGTCTATTTCATCTCTTCAAGACCTGACAAATCAAAAATATCATGTCTCAAGTTCTG 60  
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 DB 61 ACATCGACACCCAGGCGGTGCCCCCAACAAAGATTAAGCTGTCTGGCAACGAACCAACCA 120  
 QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
 DB 121 ATTCACACCAACACAGTCAAGGTAAACACACGACATGGAAGCCGATGCTATTGCTGCT 180  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
 DB 181 GCTTCTGGAAAAACAAACTTCTCTGACATACAAAAACAAACAGCTCCACAAACAGGGA 240  
 QY 84 ValAlaAlaGlySerGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
 DB 241 GTTGCTGCTGGGAAGAAATCTCAGAAAGTCAAAAGCAGAGTGTCTATCTGGAGTATCA 300  
 QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 DB 301 GAGAGCGGTGCTACTACAGCATCAAAATACGCAACAAATTTGCTATGACAGCTCTATT 360  
 QY 124 GlnGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 DB 361 GAAGAGCGGCGCAAAAGTATGAGGTACTACCTTAGAGTCACTTCAAAACCTCAGCGCGG 420  
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163

DB 421 CAAATGAAGAAGTCGAAGCGGTTGTGTGCTGCTCCCTCAGGAAAGTTCGGTTCC 480  
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183  
 DB 481 GCAAAATGTGAACACCTGAGCTCCCAAGCCGGGTGACACCAAGATGAGAGGTTATC 540  
 QY 184 GlnIleGlnLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSerAla 203  
 DB 541 GAATTCGAGCTCGGCTTGGTCAAAAGCAATTCAGACATTTGGAGAGAGCCACAAATTCGCC 600  
 QY 204 LeuSerAsnTyraLaserThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223  
 DB 601 TTATCTAAGTATGCAAGTACACAGCAGCAAGACAGCAACCAAAATAAGTCTGAGAA 660  
 QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnThrGlnGlnMetLysAlaAlaGlu 243  
 DB 661 AAGCAAGCGATTAATAAATCGATTAAGAACGAGAAACCAATACCAAGATGAAGGCTGCCGA 720  
 QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
 DB 721 CAGAACTCTAAGATCTTCGAAAGCAATGATGATCTCATCTGATGATCGCGGTT 780  
 QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
 DB 781 TCTGTCCCAATTACAGTATTATTTCTGCTCTATTCTTTTACATCGGAGCTGCACTC 840  
 QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaAlaGlyAlaAlaGly 303  
 DB 841 GCTGGAGCTCCCTGGGAGGTGCTGTAGTGACCGGCACTGGAGGTGACGAGGAGCT 900  
 QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323  
 DB 901 GCTGCCCAACACGAGTACGACACCAATTAACATTAAAGCTGTGTCACAGCGGTGAAA 960  
 QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
 DB 961 CAACTGTTTACACAGCTGCAGACAGCAAGCGATCACCGCGCTTAAAAAGCGGTGCATA 1020  
 QY 344 SerGlyLysLeuAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
 DB 1021 TCTGGAATTAAGCATTTATCAAAAACCTTAGTCAAAACGATGGCAAGCCATTCTTCTAAA 1080  
 QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
 DB 1081 GGAATCTTAAGTTTTCGCTAAGGAGACTCAATGATTTGCGAGAACTTCCCAAGCTTC 1140  
 QY 384 SerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValAlaVal 403  
 DB 1141 TCGAAAGTCAATCTGCTCTTACACAGTAATAATGGGTACGGTGGGGTTGGTGTAGTT 1200  
 QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423  
 DB 1201 GCGGCGCTCTCTCGGTAAAGGATTAATGCAATGACGCTCTCGGAGTCAACAAAC 1260  
 QY 424 ValAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
 DB 1261 GTGCTCAATTTTCGAAGAAGTGCAGAAACTCGAGGCTGCGGCTGATATGATTTTCATG 1320  
 QY 444 PheThrGlnPheThrPrgGlnAlaLysSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
 DB 1321 TTCACCTCAATTTTGGCAACAGCGCAAGTAATAATTGCTCAAAACCAACAGCGAGTCAAT 1380  
 QY 464 GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyraAlaAlaIle 483  
 DB 1381 GAATGATCTCAAAAAGCTTACCAAGCTGGCGCTCAAAATCTTAAAGGTATGCGCGCATTC 1440  
 QY 484 SerGlyValIleAlaGlyAla 490  
 DB 1441 AGCGAGACCATCGCTGGCGCA 1461  
 RESULT 13  
 E12535

Db 1501 AATGAATGACTCAAAAAGCTACCAAGCTGGCGCTCAATCTTAAAGCGTATGCCGA 1560  
Oy 483 IleserGlyAlaIleAlaGlyAla 490  
Db 1561 ATCAGCGGAGCCATCGCTGGCGCA 1584  
RESULT 11  
ARI22105  
LOCUS ARI22105 1464 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 3 from patent US 6165478.  
ACCESSION ARI22105  
VERSION ARI22105.1 GI:14106422  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 1464)  
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.  
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide  
JOURNAL Patent: US 6165478-A 3 26-DEC-2000;  
FEATURES  
source 1. 1464  
/organism="unknown"  
BASE COUNT 470 a 331 c 348 g 315 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.05e-117 Length: 1464  
Score: 2316.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
DB: Gaps: 0  
US-09-889-314-2 (1-496) x ARI22105 (1-1464)  
Oy 4 MetSerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValIleu 23  
Db 1 ATGCTAATTTCATCTTCCTCAGACCTGCACATCAAAAAATATCATGTCTCAAGTTCG 60  
Oy 24 ThrSerThrProGlnGlyAlaProGlnGlnAspIlyLeuSerGlyAsnGluThrLysGln 43  
Db 61 ACATGACACACCCAGGCGCTGCCCAACAAAGATAGCTGTGCGAAGCAAGCAAGCAA 120  
Oy 44 IlegInGlnThrArgGlnGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
Db 121 ATACGACCAACACGTCAGGGTAAACACTGAGATGGAAGCGATGCCACTATGTCTGCT 180  
Oy 64 AlaSerGlyLysAspIlySerThrSerThrThrLysThrGluThrAlaProGlnGlnGly 83  
Db 181 GCTTCGGAAGACAAAACCTTCGACTACAAAACGAAACAGAGCTCCACACAGCGA 240  
Oy 84 ValAlaIleAlaGlyLysGlnSerSerGlnIlyAlaGlyAlaAspThrGlyValSer 103  
Db 241 GTTGCTGCTGGGAAAGATCCTCAGAAAGTCAAAAGCGGTGCTGATCTGAGATATCA 300  
Oy 104 GAlaValAlaIleAlaThrThAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
Db 301 GAGGCGGCTGCTACTACACATCAATACTGCACAAAATTTGCTATGCGACCTCTATT 360  
Oy 124 GlnGlnAlaIleSerLysSerMetGlnSerThrLysGlnSerLysGlnSerLysSerAlaIle 143  
Db 361 GAAGAGCGGAGCAAAAAGTATGAGTCTACCTAGATCAAGCTCAAAAGCCTCACTGCGGG 420  
Oy 144 GlnMetLysGlnValAlaGlnAlaValAlaIleAlaLeuSerGlyLysSerSerGlySer 163  
Db 421 CAAATGAAGAAGATCGAAGCGCTGTGTGCTGCCCTCAGGGAAGGTTCCGGGTTCC 480  
Oy 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183  
Db 481 GCAAAATTTGGAACACCTGAGCTCCCAAGCCCGGGGTACACCAAGATCAGAGGTTATC 540  
Oy 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLysGlnGlnAlaThrLysSerAla 203

Db 541 GAAATCGGACCTCGCCCTTCCTAAAGCAATTCAGACATTCGGAGAACCAAAAATCTGCC 600  
Oy 204 IeuserAsnTyAlaSerThrGlnAlaGlnAlaAspGlnThrAspLysLeuGlnIleu 223  
Db 601 TTATCTTAACCTATGACAGATACACAGCACACAGCAAGCAAAATTAATCTAGTCTTAGAA 660  
Oy 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIlyThrGlnIleMetLysAlaIleGln 243  
Db 661 AAGCAAGCATATAAAATCATTAAGACAGACAGAAAGATACCAAGACATGAAGCTCCGCA 720  
Oy 244 GlnLysSerLysAspLeuGlnIlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
Db 721 CAGAAGCTAAAGATCTCGAAGGAACAATGATACTGTCATATCTGATGATCCGGGTT 780  
Oy 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
Db 781 TCTGTTCCTATTACAGTATATTTCTATGTTCGCTATTTTACATGCGGAGCTTGACTC 840  
Oy 284 AlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303  
Db 841 GCTGGACTCGCTCGCGGAGCTGCTGAGTGCAGCGGCGAGCTGAGTCCACAGAGCT 900  
Oy 304 AlaAlaIleAlaThrThrValAlaIleThrGlnIleThrValGlnAlaValAlaValLys 323  
Db 901 GCTGCCGCAACACGAGTACACACAAATTAACAGTTCAAGCTGTGTCCAAACGGTGAA 960  
Oy 324 GlnAlaValAlleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
Db 961 CAAGCTGTTATACAGCTGTCAGACAGGATCACCGCGGCTATAAACGGGCTGTCAAA 1020  
Oy 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
Db 1021 TCTGGATATAAGCATTTTTCAAAACCTTAGTCAAAAGCATTCGCAAGCCATTCTTAA 1080  
Oy 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
Db 1081 GGAATCTTAAGGTTTTCGCTAAGGAACTCAATGATTCGCAAGAACTTCCCAAGCTC 1140  
Oy 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403  
Db 1141 TCGAAAGTATCTCTCTCTTACAGTAATAGGTCACAGGTGGGGTGGGTGTGATTT 1200  
Oy 404 AlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnIleAsn 423  
Db 1201 GCGGCGCTGCTCTCGGTAAGGATTAATGCAAAATGCACCTCCGAGATGCCAAGAAC 1260  
Oy 424 ValAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
Db 1261 GTGCTCAATTCAGAAAGACGCGAANAACAGAGGCTGCGGTGATGATTTCTATG 1320  
Oy 444 PheThrGlnPheThrGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
Db 1321 TTCACCTCAATTTTGGCAACAGGCAAGTAAATTTGCTCAAAAACAAACAGCGCACTTAAT 1380  
Oy 464 GlnMetThrGlnLysAlaThrLysLeuGlnAlaGlnIleLeuLysAlaIleAlaIle 483  
Db 1381 GAAATGACTCAAAAAGCTACCAAGCTGGGCGCTCAATCTTAAAGCGTATGCCCAATC 1440  
Oy 484 SerGlyAlaIleAlaGlyAla 490  
Db 1441 ACGGAGCCATCGCTGGCGCA 1461  
RESULT 12  
E12520  
LOCUS E12520 1464 bp DNA linear PAT 27-APR-1998  
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.  
ACCESSION E12520  
VERSION E12520.1 GI:3251353  
KEYWORDS JP 1997009974-A/L.  
SOURCE unidentified.  
ORGANISM unidentified.  
unclassified.



polypeptide fused protein.  
ACCESSION E16671  
VERSION E16671.1 GI:5711354  
KEYWORDS JP 1998212298-A/1.  
SOURCE unidentified.  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1587)  
AUTHORS Izuetsu,H. and Hagimwara,T.  
TITLE HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY  
Patent: JP 1998212298-A 1 11-AUG-1998;  
JOURNAL HITACHI CHEM CO LTD  
COMMENT OS None  
OC Artificial sequences.  
PN JP 1998212298-A/1  
PD 11-AUG-1998  
PF 31-JAN-1997 JP 1997018522  
PI IZUTSU HIROSHI, HAGIMWARA TOSHIKATSU  
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC C12P21/08,  
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC C12R1:91),  
PC (C12P21/02,C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key  
FH Location/Qualifiers  
FT source 1..1587  
FT Location/Qualifiers  
FT 1..1587  
/organism='Artificial sequences'  
source  
1..1587  
/organism='unidentified'  
/db\_xref='taxon:3264'  
BASE COUNT 499 a 363 c 384 g 341 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,94e-117 Length: 1587  
Score: 2317.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.80% Mismatches: 0  
Query Match: 97.80% Indels: 0  
Gaps: 0  
US-09-889-314-2 (1-496) x E16671 (1-1587)  
QY 3 AsnMetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22  
Db 121 TCATGTCATTTTCATCTCTTCAGACCTGACAAATCAAAAAAATATCATGTCTCAAGTT 180  
QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLys 42  
Db 181 CTGACTCGACACCCAGGCGCTGCCCAACAAGATTAAGTGTCTGGCAACGAAACGAG 240  
QY 43 GlnIleGlnIleThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAla 62  
Db 241 CAAATACAGCAACACGCTCAGGTAATAAACACTGAGATGGAAGCCGATGCCACTATTGCT 300  
QY 63 GlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGln 82  
Db 301 GGGCGCTTGGAAAAAGCAAACTCTCGACTACAAAAAAGAAAGAGAGCTCCACAAACAG 360  
QY 83 GlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyVal 102  
Db 361 GGGATTGCTGCTGGAAAGATCTCTGAAAGTCAAAGGACAGGTGCTGATTACGAGTA 420  
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122

|||||  
Db 421 TCAGAGACGGCTGCTACTACAGCATCAATATCTGCACAAAAATTGCTATGACAGACTCT 480  
QY 123 IlleGlnLysAlaSerLysSerMetGlnSerThrLeuLysLeuGlnSerIleSerAla 142  
Db 481 ATTGAAAGGCGAGCAAAATATGAGTCTTACCTTAAGTCACTTCAAACTCAGTGGCC 540  
QY 143 AlaGlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGly 162  
Db 541 GCGCAATATGAAGAGTCGAAGCGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600  
QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlyVal 182  
Db 601 TCCGCAAAATTTGGAAACACCTGACCTCCCAAGCCGGGGGACACCAAGATCAGAGCTT 660  
QY 183 IlleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSer 202  
Db 661 ATCGAAATCGGACCTCGGCTGCTTGAAGCAATTGACATTTGGAGAGAGCCACAAAAATCT 720  
QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222  
Db 721 GCCTTATCTACTATGCAAGTACAAAGCAAGACAGACCAACAAATTAACCTAGGTCTA 780  
QY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysThrGlnLysMetLysAlaAla 242  
Db 781 GAAAGACAGACGATAAATAATCGATTAAGAAAGCAAGAAATACCAAGATGAAAGCTGCC 840  
QY 243 GlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262  
Db 841 GAACAGAGAGCTTAAAGATCTCGAAGGAAACAATGATCTCTCAATCTCGATGATGATCGCG 900  
QY 263 ValSerValAlaIleThrThrValIleSerIleValAlaAlaIlePheThrGlyGlyAlaGly 282  
Db 901 GTTCTGTGCTACTTACAGTATATTCTATTGTGTCTCTATTATTATCGGGAGCTGGA 960  
QY 283 LeuAlaGlyLeuAlaAlaGlnAlaAlaValAlaGlyAlaAlaAlaGlyAlaAlaGly 302  
Db 961 CTGCTGTGACTCGCTCGGGAGCTGCTGTAGTGCACAGCGGACGCTGGAGTGCACAGCGA 1020  
QY 303 AlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322  
Db 1021 GCTGCTGCCCGCAACACGAGTACCAACAATTCACGTTCAAGCTGTGTTGCCAAGCGCTG 1080  
QY 323 LysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342  
Db 1081 AAACAGCTGTATACAGCTGTACAGACGATCCACCGCGGTATTAAGGCGCTGC 1140  
QY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362  
Db 1141 AAATCTGCAATTAAGCATTTATCAAAAACCTTAGTCAAAAGCGATTCGCAAAAGCATTTCT 1200  
QY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382  
Db 1201 AAAGGAATCTCTTAAGGTTTCGTTAAGGAATCAATTAATGGAAGAACTTCCCAAG 1260  
QY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValAl 402  
Db 1261 CTCTCGAAATCACTCTGCTCTTACCAAGTAAGAGGGTCAAGGTGGGTGGGTTGTA 1320  
QY 403 ValAlaAlaProAlaLeuGlnGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 422  
Db 1321 GTTCCGGCGCTGCTCTCGGTAAAGGATATATGAAATGACAGCTCTCGAGATGCAACAA 1380  
QY 423 AsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442  
Db 1381 AACGTGCTCAATTTTCAGAAAGAGTCGGAATACTGCAAGCTGGCGCTGATATGATTTCT 1440  
QY 443 MetPheThrGlnPheThrProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSer 462  
Db 1441 ATGTTCACTCAATTTTGGCAACAGCAAGTAATTTCCCTCAAAAACAAAGCGGAGTCT 1500  
QY 463 AsnGlnMetThrGlnLysAlaThrLysLeuGlyValAlaGlnIleLeuLysAlaThrAlaAla 482  
|||||



PN JP 1998212298-A/7  
 PD 11-AUG-1998  
 PF 31-JAN-1997 JP 1997018522  
 PI IZUNSU HIROSHI, HAKIMARA TOSHIKATSU  
 PC C07K16/12, C07K19/00, C12N5/10, C12N7/00, C12M5/09, C12P21/02, PC  
 C12P21/08,  
 PC G01N33/531, G01N33/571//A61K39/395, C07K14/295, (C12N5/10, PC  
 C12R1.91),  
 PC (C12P21/02, C12R1.91):  
 CC strandedness: Double;  
 CC topology: Circular;  
 FH Location/Qualifiers  
 FT source 1..11099  
 FT FT /organism='Artificial sequences' FT CDS  
 FT 4135..5601 /product='Chlamydia pneumoniae antigen FT  
 FT misc\_feature /polypeptide'  
 FT 7..4029  
 FT /note='DNA sequences derived from FT  
 FT baculovirus' 5635..8467  
 FT misc\_feature /note='DNA sequences derived from FT  
 FT baculovirus' 4030..4092.  
 FT promoter  
 FT Location/Qualifiers  
 1..11099  
 /organism='unidentified'  
 /db\_xref='taxon:32644'  
 BASE COUNT 3075 a 2453 c 2525 g 3046 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,81e-116 Length: 11099  
 Score: 2320.00 Matches: 490  
 Percent Similarity: 99.60% Conservative: 2  
 Best Local Similarity: 99.19% Mismatches: 1  
 Query Match: 97.93% Indels: 1  
 Gaps: 0

US-09-889-314-2 (1-496) x E16677 (1-11099)

QY 3 AsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22  
 DB ::  
 DB 4132 TCcATGTCATATTCATCTCTTCAGACCTCGACATCAAAAATATTCATGCTCAGTT 4191  
 QY 23 LeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGlnThrLys 42  
 DB ::  
 DB 4192 CTGACATCGACACCCCGAGGCGTGCCTCCCAAGATTAAGCTGTCTGGCAACGAAG 4251  
 QY 43 GlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspIleThrIleAla 62  
 DB ::  
 DB 4252 CAAATACAGCAACACGTGAGGTAAACACTGACGTGGAAGCGATGCTACTATTGCT 4311  
 QY 63 GlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGln 82  
 DB ::  
 DB 4312 GGTGCTTCGGAAGAACAAACTCTCTCGACACACAAAAACAGAACAGCTCCAAACAG 4371  
 QY 83 GlyValAlaAlaGlnLysGlnSerSerGlnSerGlnLysAlaGlnAlaAspThrGlyVal 102  
 DB ::  
 DB 4372 GGAGTGTGCTGCTGGGAAGATCTCGAAGACCAAAAGCAGAGTGTCTACTGAGTA 4431  
 QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122  
 DB ::  
 DB 4432 TCAGGAGCGGCTCTACTACATCAAAATACGCAACAAAATGCTATGACAGACCTCT 4491  
 QY 123 IleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla 142  
 DB ::  
 DB 4492 ATTGAAGAGCGGAGCAAAATATGAGTCTAGAGTACCTTCAACACCTCAAGGCC 4551  
 QY 143 AlaGlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGly 162  
 DB ::  
 DB 4552 GCCCAATATGAAGAGTCGAAGCGGTTGTGTGCTGCCCTTCAGGGAAGATTCGGGC 4611

QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182  
 DB ::  
 DB 4612 TCcGCAAAATTCGAACACCTAGCTCCCAAGCCGGGGTACACCAAGATCAAGGTT 4671  
 QY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSer 202  
 DB ::  
 DB 4672 ATCGAAATCGGAGCTCGCGCTTCTTAAGCAATTCAGACATTGGGAGAGCCCAAAATCT 4731  
 QY 203 AlaLeuSerAspThrAlaSerThrGlnGlnAlaAspGlnThrAsnLysLeuGlnLeu 222  
 DB ::  
 DB 4732 GCCATTATCTAATGACAGTACACAGCAAGCAAGACCAACAAATTAATTAAGTCTA 4791  
 QY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysGlnLysAla 242  
 DB ::  
 DB 4792 GAAAGCAAGCATATAAATCATTAAGACAGAGAAGATACCAAGATGAAGGCTGCC 4851  
 QY 243 GlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262  
 DB ::  
 DB 4852 GAACAGAACTTAAGATCTCGAAGAACAAATGATACTGTCAATATCTGATGATCGCG 4911  
 QY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGly 282  
 DB ::  
 DB 4912 GTTCTGTGTCATTAACATTAATTTCTATTGTGCTGCTATTATTACATCGGAGCTGGA 4971  
 QY 283 LeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGly 302  
 DB ::  
 DB 4972 CTCGCGAGACTGCTCGGAGACTGCTGTAGGTGCGAGCGGAGCTGAGGTGCGAGCAGA 5031  
 QY 303 AlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322  
 DB ::  
 DB 5032 GCTGTCGCCCAACACCGGTAGCAACAAATTAAGTTCAGTTCAGTTCAGTTCAGGCGTG 5091  
 QY 323 LysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342  
 DB ::  
 DB 5092 AAACAAGCGTTATATCACCTGTGACAGCAAGCATACCGCGCTTTAAAGGGCTGTC 5151  
 QY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362  
 DB ::  
 DB 5152 AATCTGGAATAAACATTTATCAAAACTTATAGTCAAAAGCGATTGCCAAAGCATTTCT 5211  
 QY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382  
 DB ::  
 DB 5212 AAAGCATCTCTTAAGCTTTTCGCTAAGGACCTCAATATTCGGAAGACTTCCCAAG 5271  
 QY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyAlaVal 402  
 DB ::  
 DB 5272 CTCTCGAAAGTCATCTGCTCTTACCAAGTAAATGGGTCAAGGTGGGTTGGGTTGTA 5331  
 QY 403 ValAlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 422  
 DB ::  
 DB 5332 GTTGGCGGCGCTCTCTCGTAAAGGATTATGCAATGACAGCTTCGAGATGCAACAA 5391  
 QY 423 AsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442  
 DB ::  
 DB 5392 AACGTGCTCAATTTGCAAAAGAGTCGGAACAACTCAGGCTGCGCTGTATGATTTCT 5451  
 QY 443 MetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSer 462  
 DB ::  
 DB 5452 ATGTTCACTCAATTTTGGCAACAGCAAGTAATATGCTCAAAACAAACAGCGAGTCT 5511  
 QY 463 AsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAla 482  
 DB ::  
 DB 5512 AATGAATGACTCAAAAGAGTACCAAGCTGGGCGCTCAAAATCTTTAAAGGATGTCGCA 5571  
 QY 483 IleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsn 495  
 DB ::  
 DB 5572 ATCAGCGGAGCATTCCTGCGCAGCATTAAGCTTAAGT 5611  
 RESULT 10  
 E16671  
 LOCUS E16671 1587 bp DNA linear part 28-JUL-1999  
 DEFINITION DNA encoding histidine tag - Chlamydia pneumoniae antigen

```

promoter          4030..4092
FT               /note='Baculovirus polyhedrin gene promoter'
FT CDS           4135..5601
FT              /product='Chlamydia pneumoniae 53kDa-antigen
FT              peptide'
FT              5635..8467
FT              /note='Baculovirus-derived sequence'.
FEATURES
  source          1..11099
                  /organism='unidentified'
                  /db_xref='taxon:32644'
BASE COUNT      3075 a 2453 c 2525 g 3046 t
ORIGIN
Alignment Scores:
Pred. No.:      2,81e-116      Length:      11099
Score:          2320.00      Matches:      490
Percent Similarity: 99.60%      Conservative: 2
Best Local Similarity: 99.19%      Mismatches: 1
Query Match:    97.93%      Indels:      1
DB:             6      Gaps:      0
US-09-889-314-2 (1-496) x E16636 (1-11099)

QY 3 AsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22
Db 4132 TCATGCTCATATTTCATCTTCTTCAGACCTGCATCAATAAATATCATGCTCAAGTT 4191

QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLys 42
Db 4192 CTGACATGACACCCAGGCGGCTGCCACACAGATAGCTGTGGCAGACGAAAGCAAG 4251

QY 43 GlnIleGlnInThrArgGlnGlyLysAsnThrGlnMetGlnLysSerAspAlaThrIleLea 62
Db 4252 CAANACAGCAACACGTCAGGCTAAACACCTGATGAGAAAGCGATGCGCATATTGCT 4311

QY 63 GlyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaProGlnGln 82
Db 4312 GGTGCTCTGGAAAAAGACAAACTCCCTCGACTACAAAAACAGAAACAGCTCCACACAG 4371

QY 83 GlyValAlaAlaGlnGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyVal 102
Db 4372 GGAGTGTGCTGGAAAAAGATCCCTCAGAAAGTCAAAAGGAGGCGCTGATCTGAGCTA 4431

QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 4432 TCAGAGGCGGCTGCTACTACAGCATCAAAATCTGCACAAATAATTCATGACAGCTCT 4491

QY 123 IleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla 142
Db 4492 ATTGAAGAGCGGCAAAAGATAGAGTCTACCTTAGAGTCACTTCAAAAGCCTCAGTCC 4551

QY 143 AlaGlnMetLysGlnValGlnAlaValValAlaAlaAlaLeuSerGlyLysSerSerGly 162
Db 4552 GCGCAAAATGAAAGAGTCGAAAGCGGTGTTGCTGCTCCCTCAGGAAAAAGTTCCGGCC 4611

QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182
Db 4612 TCAGCAAAATGGAACACCTGAGCTCCCGGAGCGGGGAGACACCAAGATCCAGAGTT 4671

QY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSer 202
Db 4672 ATCGAAATTCGAGACTCGCGCTTGCTAAAGCAATTGACACATTGGGAGAAAGCAAAATCT 4731

QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
Db 4732 GCCCTATCTACTATGCAAGTACACAGCAAGCAACACCAAAATAAATTAAGTGGTCTA 4791

QY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnThrGlnGlnMetLysAlaAla 242
Db 4792 GAAAAAGCAAGGATAAATAATGATTAAGAACGACAAATATCCAAAGATGATGAAGGCTGCC 4851

QY 243 GlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262

```

```

Db 4852 GAAACGAGACTCTAAAGATCTCGAAGAAACAATGCTACTGTCATATTGATGATCGCG 4911
QY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGly 282
Db 4912 GTTTCGTGTGGCATATACAGTTATTTCTATGCTGTGCTGCTATTTTACATGGCGAGCTGGA 4971
QY 283 LeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGlyValAlaGly 302
Db 4972 CTCGCTGAGCTCGCTGCGGAGAGCTGCTAGCTGCAGCGGAGCGTGGAGGTGACAGCA 5031
QY 303 AlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322
Db 5032 GCTGCTGCCGCAACACCGCTGACCAACAATTCAGATTCAGCTGTTGTCCAAAGCGGTG 5091
QY 323 LysGlnAlaValAlaIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342
Db 5092 AAACAAAGCTGTATACACAGCTGTGACAGCAAGCATCACCGCGGCTATTAAGGCGGTGTC 5151
QY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362
Db 5152 AAATCTGGAATTAAGCAATTTATCAAAACTTTAGTCAAAAGCGATTTGCCAAAGCATTTCT 5211
QY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382
Db 5212 AAAGAAATCTTAAGCTTTTCCTTAAGGAACTCAAAATGATGGCAAGAACTTCCCAAG 5271
QY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValAla 402
Db 5272 CTCCTGAAAGTCATCTGCTCTTACAGATTAATGGGTACAGGTGGGGGTGGGGTGTGA 5331
QY 403 ValAlaAlaProAlaLeuGlnGlyLysGlyIleMetGlnMetGlnLysSerGlnMetGlnGln 422
Db 5332 GTTGGGCGGCTGCTCTCGGTAAGGATTAATGCAAAATGCAAGCTCTCGGAGATGCCAA 5391
QY 423 AsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442
Db 5392 AACGTCGCTCAATTTTCAGAAAGATGCGGAATACTGCAGCTGCGGCTGATATCATTTCT 5451
QY 443 MetPheThrGlnPheThrPheGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSer 462
Db 5452 ATGTTCACTCAATTTTGGCAACAGCAAGTAATTAATTCCTCAAAACAAACAGCGAGTCT 5511
QY 463 AsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAla 482
Db 5512 AATGAATGACTCAAAAGACTACCAAGCTGGCGCTCAAAATCTTAAGCGTATGCGCA 5571
QY 483 IleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsn 495
Db 5572 ATCAGCGGAGCCATCGCTGGCGGCAAGCATTAAGCTAGAAAT 5611

RESULT 9
LOCUS      E16677
DEFINITION 11099 bp DNA linear PAT 28-JUL-1999
            antigen polypeptide-coding DNA.
ACCESSION  E16677
VERSION    E16677.1 GI:5711360
KEYWORDS   JP 1998212298 -A/7.
SOURCE     unidentified.
ORGANISM   unidentified.
REFERENCE  1 (bases 1 to 11099)
AUTHORS   Izutsu,H. and Hagiwara,T.
TITLE     HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION
          PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN,
          RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE
          RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND
          PRODUCTION OF ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY
          Patent: JP 1998212298-A 7 II-AUG-1998;
          HITACHI CHEM CO LTD
COMMENT    OS None
          OC Artificial sequences.

```



```

|||||
Db 16236 ATTGCTGCTCTTGGAAAAAGACAAACTTCCGCTGCTCAAAAACAGAAAGCGCTCA 16177
QY 81 GlnGlnGlnValAlaAlaGlyLysSerSerGlnSerGlnLysAlaGlyAlaAspThr 100
Db 16176 CAACAGGAGTCTCTGCTGGAAAAAATCTCTGAAAAAGGCAAAAGGCGCTGCAATCT 16117
QY 101 GlnValSerGlnValAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
Db 16116 GAGGTATCAGAGAGCGGTGCTACTACAGCATCAAAATCTGCAACAAAAATTCCTATG 16057
QY 121 ThrSerIleGlnLysAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
Db 16056 ACCTTATTAAGAGGCGACCAAAAGTAGTAGTCTACCTTAGAGTCACTTCAAAAGCTC 15997
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSer 160
Db 15996 AGTGCCTGCGCAAAATGAAAGAGTCGAAGCGGTGTTGCTGCTGCTCAGGGAAGT 15937
QY 161 SerGlnSerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSer 180
Db 15936 TCGGGTTCCTCCAAATTTGGAAACACTGAGCTCCCAAGCCGGGGTGACACCAAGATCA 15877
QY 181 GlnValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysAlaThr 200
Db 15876 GAGGTATCGAAATCGGAGCTCGGCTTGCTAAAGCAATTCAGACATTTGGAGAACCCACA 15817
QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 15816 AAATTCGCTTATCTACTACTATGACATGACAAAGCACAAACAGCCCAAAACAAATTAACCTA 15757
QY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGln 240
Db 15756 GGTCTAGAAAAAGACAGCGATTAATAATCGATTAAGAACGAGCAATACCAAGAGATGAG 15697
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db 15696 GCTGCCGAACAGAAAGTCTTAAGATCTGAAAGAAACATGATCTCACTACTGCTGATG 15637
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280
Db 15636 ATCGCGGTTTCTGTCGCAATACAGTTATTCATTTCTGCTCTATTTTAACTGGCGGA 15577
QY 281 AlaGlnLysLeuAlaGlnLysLeuAlaAlaGlyAlaAlaValGlnAlaAlaAlaGlyAla 300
Db 15576 GCTGGAGCTCGCTGACTCGCTGCGGAGCTGCTGTAGCTCAGCGGAGCTGGAGGTGCA 15517
QY 301 AlaGlnAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln 320
Db 15516 GCAGGAGCTGCTCGCCGCAACACGAGTAGCAACAAATTTACAGTTCAGCTGTGTCCAA 15457
QY 321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
Db 15456 GCGGTAAACCAACCTCTTATCACAGCTGTGACAAAGCGATCACCGGCTATAAAACGC 15397
QY 341 AlaValLysSerGlyLysLeuAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
Db 15396 GCTGTAAATCTGGATTAAGCATTTATCAAACTTTACTCAAGCGATTGGCAAGGCC 15337
QY 361 IleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
Db 15336 ATTTCTAAAGAGATCTCTAAGTTTCGTAAGGAACTCAAAATGATTCGGAAGACTTC 15277
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGly 400
Db 15276 CCCAAAGCTCTCGAAAGTCATCTGCTCTTACACAGTAAATGGGTTCACGGTTGGGG 15217
QY 401 ValValValAlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMet 420
Db 15216 GTTGTAGTTGGCGGCTGCTGCGTAAAGGATTAATGCAATTCAGCTCGAGATG 15157
QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlnValGlnLysLeuGlnAlaAlaAspMet 440
|||||

```

```

Db 15156 CAACAAACGTCGCTCAATTTAGAAAGAGTCGAAACACTGCAAGCTGCGCTGATATG 15097
QY 441 IleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460
Db 15096 ATTTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCTCTCAAAACAAACAGC 15037
QY 461 GlnSerAsnGlnMetThrGlnLysAlaThrLysLeuAlaGlnIleLeuLysAlaThr 480
Db 15036 GAGCTAATGAAATGACTCAAAAAGCTACCAAGCTGGGCGCTCAAAATCTTAAAGCTAT 14977
QY 481 AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsnPhe 496
Db 14976 GCGCAATCAGCGGAGCGATCGCTGGCGAGCATAAACCAATTAATTTT 14928
RESULT 7
E16678
LOCUS
DEFINITION
All sequences of plasmid pHTCPN53 containing DNA encoding histidin
tag - Chlamydia pneumoniae antigen polypeptide fused protein.
ACCESSION
E16678.1 GI:5711361
VERSION
JP 1998212298-A/8.
KEYWORDS
unidentified.
SOURCE
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 9595)
AUTHORS
Izutsu,H. and Hagiwara,T
TITLE
HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION
PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN,
RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE
RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND
PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY
Patent: JP 1998212298-A 8 11-AUG-1998;
JOURNAL
HITACHI CHEM CO LTD
COMMENT
OS None
OC Artificial sequences.
PN JP 1998212298-A/8
PD 11-AUG-1998
PE 31-JAN-1997 JP 1997018522
PI IZUTSU HIROSHI, HAGIWARA YOSHIKATSU
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC
C12P21/08,
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC
C12R1:91),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Circular;
CC hypothetical: No;
CC anti-sense: No;
FH key
FH Location/Qualifiers
FT source 1..9595
FT /organism='Artificial sequences' FT CDS
FT 2206..3795
FT /product='Histidin tag - Chlamydia pneumoniae
antigen
polypeptide fused protein'
FT misc_feature 1..2128
FT /note='DNA sequences derived from FT
baculovirus'
FT misc_feature 3868..6480
FT /note='DNA sequences derived from FT
baculovirus'
FT 2129..2191.
FT promoter
FT Location/Qualifiers
FT 1..9595
FT /organism='unidentified'
FT /db_xref='taxon:32644'
FEATURES
source
BASE COUNT 2676 a 2114 c 2191 g 2614 t
ORIGIN
Alignment Scores: 1.12e-116 length: 9595
Pred. No.:

```

```
gene
CDS
LDD"
/complement(2558. .3607)
/gene="CPj0799"
/complement(2558. .3607)
/gene="CPj0799"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA99007.1"
/db_xref="GI:8979173"
/translation="MAAIKQILRSMLOSSSLMMVLESLYSLSGTCYITTDKPEDFHS
SSAVKMDHMKTTLSRLSNKRAKAAVSGTATVVGFTKDTWSRTYAVRWYWGKEL
PTSSWVKSKATGSSDGSIIAGIYENELSSQFAVTMKNEMYLPTWAVQSKAYGI
SSDSGLYVGSADAMSRFEAVKMGHEAOVLPGVAVSANSVANSANGSIYVGSODA
SGILYAVKMGENTITHTGLTGYSALAAKAVNNGVIYGRSETTGYEPFHACHKNGVM
SDLTIGGSIYSAKGVSAATGKVIYGMSTTANGKLHAFYVCGRMIDLEIYSWKECAN
AVSIDGETIVGQSE"
/complement(3867. .5153)
/gene="eno"
/complement(3867. .5153)
/gene="eno"
/codon_start=1
/transl_table=11
/product="enolase"
/protein_id="BAA99008.1"
/db_xref="GI:8979174"
/translation="MFEAVIADIQARELIDSRGPTLVKVTSTSGVGEARVPSGAS
IGKKEALEFRDTSFRIOGKVLQAVKVKETLEPVLVGCYSYEDSLIDSLMDSGDS
PNKETLGNAILIGVSLATAHAAATLRPLRYLGGCFACSLPCPMMLINGCMADN
GLGQEFMRIRIGASIKAEVNMGADVHTLKLHERGLSTGVDEGGFAPNLASNE
EALDELILAIEKAGFTPGKDISLADCAASFEYNKGTGDRHYEEOIALISNCDR
YPISDIEDGLAEEDYDGMALLTEVLGEKVOIVGDDLFTVNPDELILEGINSGLANVLI
KPNQIGLTETEVVAIKLAOMAGYTTIISHRSETTDTIADLAFAFNAGQIKTGSLSR
SERAKVYRMLEIEELGSEAFITDSNVFTEDSEE"
/complement(5345. .7318)
/gene="uvrB"
/complement(5345. .7318)
/gene="uvrB"
/codon_start=1
/transl_table=11
/product="exonuclease ABC subunit B"
/protein_id="BAA99009.1"
/db_xref="GI:8979175"
/translation="MTFOLHAPFAPCGDPPAIALRSAGVRNOVKSQVLLGTGSGKT
FTIANVYANVLPTLVLAHNKTLAOLYQEFREPPNNAVEYFISYVDYQPEAYIAR
SDTIEKSLILNDELIDKLRLSATRSILERRDTLYSVSCITGTSPPNTYSMAVLE
VGKEYPRNITLAQLVKMYQASPLPQSAFERGSDVIDIFPAYESELALRLLEFLNDTL
TSIESDPLTMIIPKESVPSATLYPGSHYVPEAIRQOARTIOELEERMAFFDRPI
EKDRIFHTTIDIMEIKEIFGCKGIENYSRHTGAPPGAPPTCLLDYPPEDFLITDE
SHOTLPOIRAMYRGDSRGSILVEYGFRLPSAFDNRPLTYEAKQYFRKVIYVATPG
DVEYESGHTVVOOIRPTGTIPDMPETRPATGVDLLEETRLRLSKRHKIIVIST
TKKLAEDBAGFLSELETPAATLHSGIETAEKTOIITLDRSGVIDVLIGVNLREGLDL
PEVSLVALIDADKDEFLKSTSLIQFCGRARNINGKVIFFADQKTRSIIEETLRGTER
RROLDNKNKHNIYPPKPIKAIFANPLIQTSSKSESPKESRPSKEDLEBOIKKYE
ALMORAAKEFFNEAKYRDAMQACKBOLLVLF"
/complement(7327. .8361)
/gene="trpS"
/complement(7327. .8361)
/gene="trpS"
/codon_start=1
/transl_table=11
/product="tryptophanyl tRNA synthetase"
/protein_id="BAA99010.1"
/db_xref="GI:8979176"
/translation="MNKKRVLTGDRPPGKHLHGWSIKNRELELONSPEYDCFFII
ADLHTLTTKIRKEEVLVDNHIYEVLAOWLVSIGDPTSIITLYLOSAIPEIYELHLFS
MLISINRWKIPSLKDMARNASIEGSGISYGLIPIQSADILAKQFYPVGDNE
AHVELTRDIARNFNFLYGOVPEPEVLQGLSTLYGIGOGKMSKANNALYLSQSDA
TITEKVRKMYDNPRIKATPTPGVAGCNLFYIHDFNHKDEVEEFKAROGCICKDI
EVKARIAELIHLFLPIKERSESEPLSKLALQNLVLEDGTHKMKREYAKATMEEVHDKG
FSHKWRSLK"
/complement(8370. .8924)
/gene="CPj0803"
gene
CDS
LDD"
/complement(8370. .8924)
/gene="CPj0803"
/codon_start=1
/transl_table=11
/product="CT584 hypothetical protein"
/protein_id="BAA99011.1"
/db_xref="GI:8979177"
/translation="MAAKTKULEEDNVFELLEGNLRKIRFATPIGYTTFREONVFN
CANGQOELANFEFEELINGKLTQELAPQOKOASLIRFPMPIRVADQIHGREFIN
FTSDMLTQGERCIFLNLARVDGGEFLMTDVQNTCHLIHLLARLLEAKNPVGEK
NLQETQEEITSLKHNFDLTKALO"
/complement(8928. .9680)
/gene="gp6D"
/complement(8928. .9680)
/gene="gp6D"
/codon_start=1
/transl_table=11
/product="ChlYr plasmid paralog"
/protein_id="BAA99012.1"
/db_xref="GI:8979178"
/translation="MGNKTLLIESFRKKNPPTKKEALARKMEGDPSPILAVLSNPTL
SSKEQRLHLQHTNFRQILEBDLTQCLSLAEVKOIHQSVLHGERITTKRDL
KSYREGAFSSWLLLYGNRQTPYNFLVYELFTLPEPLKLEMEKMPQAVYTLASRQ
GPOEKKEEIIIRNYRGEKSELDRIKFEPLVETDCRKTSPVKQALMLTRGSQILTK
CTLSDSDOIILEKLIKLEKYSMLPFDTKV"
/complement(9685. .10452)
/gene="mind"
/complement(9685. .10452)
/gene="mind"
/codon_start=1
/transl_table=11
/product="chromosome partitioning ATPase"
/protein_id="BAA99013.1"
/db_xref="GI:8979179"
/translation="MKTIAVNSFKGTAKTSTTHLGAALAOYHOARVLLIDPDAQAN
LTSGLDPCDYDSLAVVLOGEKEIOEYIRPIODQDLIPADPWLRIEIVSGNLAD
RYSHERLKYVLSGVDKDYVIDPPLCLMTESALIAAYALALICATPEVSYGLE
RLAGTIGISIRARPLITLIGVALSEFWNCGKNSAFAEILHHTFPFKLLNTRIDITV
SEAHGKPYRATSPSARASDEYFWLRELLLRDI"
/complement(10485. .12392)
/gene="lhrs"
/complement(10485. .12392)
/gene="lhrs"
/codon_start=1
/transl_table=11
/product="threonyl tRNA synthetase"
/protein_id="BAA99014.1"
/db_xref="GI:8979180"
/translation="MIQVTCQKNVEVLEGTAAELAKQLNKSHQFIGVLINERPDRL
```

```
CDS
/complement(8370. .8924)
/gene="CPj0803"
/codon_start=1
/transl_table=11
/product="CT584 hypothetical protein"
/protein_id="BAA99011.1"
/db_xref="GI:8979177"
/translation="MAAKTKULEEDNVFELLEGNLRKIRFATPIGYTTFREONVFN
CANGQOELANFEFEELINGKLTQELAPQOKOASLIRFPMPIRVADQIHGREFIN
FTSDMLTQGERCIFLNLARVDGGEFLMTDVQNTCHLIHLLARLLEAKNPVGEK
NLQETQEEITSLKHNFDLTKALO"
/complement(8928. .9680)
/gene="gp6D"
/complement(8928. .9680)
/gene="gp6D"
/codon_start=1
/transl_table=11
/product="ChlYr plasmid paralog"
/protein_id="BAA99012.1"
/db_xref="GI:8979178"
/translation="MGNKTLLIESFRKKNPPTKKEALARKMEGDPSPILAVLSNPTL
SSKEQRLHLQHTNFRQILEBDLTQCLSLAEVKOIHQSVLHGERITTKRDL
KSYREGAFSSWLLLYGNRQTPYNFLVYELFTLPEPLKLEMEKMPQAVYTLASRQ
GPOEKKEEIIIRNYRGEKSELDRIKFEPLVETDCRKTSPVKQALMLTRGSQILTK
CTLSDSDOIILEKLIKLEKYSMLPFDTKV"
/complement(9685. .10452)
/gene="mind"
/complement(9685. .10452)
/gene="mind"
/codon_start=1
/transl_table=11
/product="chromosome partitioning ATPase"
/protein_id="BAA99013.1"
/db_xref="GI:8979179"
/translation="MKTIAVNSFKGTAKTSTTHLGAALAOYHOARVLLIDPDAQAN
LTSGLDPCDYDSLAVVLOGEKEIOEYIRPIODQDLIPADPWLRIEIVSGNLAD
RYSHERLKYVLSGVDKDYVIDPPLCLMTESALIAAYALALICATPEVSYGLE
RLAGTIGISIRARPLITLIGVALSEFWNCGKNSAFAEILHHTFPFKLLNTRIDITV
SEAHGKPYRATSPSARASDEYFWLRELLLRDI"
/complement(10485. .12392)
/gene="lhrs"
/complement(10485. .12392)
/gene="lhrs"
/codon_start=1
/transl_table=11
/product="threonyl tRNA synthetase"
/protein_id="BAA99014.1"
/db_xref="GI:8979180"
/translation="MIQVTCQKNVEVLEGTAAELAKQLNKSHQFIGVLINERPDRL

Alignment Scores:
Pred. No.: 1,36e-115 Length: 325865
Score: 2340.00 Matches: 496
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 98.78% Indels: 1
DB: 1 Gaps: 0

US-09-889-314-2 (1-496) x AP002548 (1-325865)
QY 1 AAPTthrAsmMetSerIleSerSerSerGlyProAspAnglnIlysAsnIleMetSer 20
|||||
DB 16416 GATACAAACATGTCTATTTTCATCTTCTTACAGACCGACAAATCAAAAAATATCATGCTT 16357
|||||
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
|||||
DB 16356 CAAGTCTGACATCGACACCCAGGCGCGCCCAACAAGATAAGCTGCTGCGAACAGAA 16297
|||||
QY 41 ThrLysGlnIleGlnGlnIleThrArgGlnGlyLysAsnThrCluMetGluSerAspIleThr 60
|||||
DB 16296 ACGAAGCAATATCAGCAACACGTCAGGGTAAACACAGTGAAGGCAATGCGACT 16237
|||||
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
|||||
```

Db 4744 TCGGGTTCGCAAAATGTGAACACACTGAGCTCCCAAGCCGGGGTGCACCAAGATCA 4803  
 QY 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200  
 Db 4804 GAGGTATTCGAATCGGACATCGCGCTTGCTAAACCAATTGACAGATTGGGAGAGCCACA 4863  
 QY 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
 Db 4864 AAMTTCCTTACTATCTACTATGCAGATACACAGACACAGACAGCCAAACAAATAACTA 4923  
 QY 221 GlyLeuGlyLysGlnAlaIleLysIleAspLysGluArgGluTyrGlnGluMetLys 240  
 Db 4924 GGCTTAGAAAAAGCAAGCGATAAAAATGATTAAGCAACGAGCAAGATACCAAGAGATGAG 4983  
 QY 241 AlaAlaGluGlnLysSerLysAspLeuGlyThrMetAspThrValAsnThrValMet 260  
 Db 4984 GCTGCCCAAGAGAGTCTAAGATCTGAAGGAACAATGATGCTCAATACGTATG 5043  
 QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280  
 Db 5044 ATCGCGTTTCTGTGTCATTCACAGTTATTCTATTGCTGCTATTATTACATGCGGA 5103  
 QY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyGlyAla 300  
 Db 5104 GCTGAGCTCCGTGACTCGCTGCGGAGAGCTGCTAGTGCAGCGGAGCTGAGGTGCA 5163  
 QY 301 AlaGlyValAlaAlaAlaIleThrValAlaIleThrGlnIleThrValGlnAlaValGln 320  
 Db 5164 GCAGGACTCTCGCCGACACCGAGTGCACAAATTCAGATTCAAGCTGTGTCGCA 5223  
 QY 321 AlaValLysGlnAlaValIleThrValAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340  
 Db 5224 CGCGTGAACACAGCTGTATTCACAGCTGTGCACAGACAGATACCGCGGCTATTAAGCC 5283  
 QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
 Db 5284 GCGTCAATCTGGAATTAAGCATTTATCAAACTTTAGTCAAAAGGATGCCAAAGCC 5343  
 QY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380  
 Db 5344 ATTTCTAAAGAACTCTTAAGTTTTCGTAAGGAACTCAAAATGCGAAGACTTC 5403  
 QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGly 400  
 Db 5404 CCCAAGCTCTCGAAGCATCTGCTCTCTTAACCGTAATGGGTACAGGTTGGGTTGGG 5463  
 QY 401 ValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMet 420  
 Db 5464 GTTGTACTTGGGGGCTGCTGCTCGGTAAAGGATATTGCAAAATGCAAGCTCTCGAGATG 5523  
 QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlyValGlyLysLeuGlnAlaAlaAspMet 440  
 Db 5524 CAACAAAACGCTCGCTCAATTTCAGAAAAGAGTCCGAAAACGTCGCGCTGATATG 5583  
 QY 441 IleSerMetPheTrpGlnPheTrpGlnAlaSerLysIleAlaSerLysGlnThrGly 460  
 Db 5584 ATTTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCCCTCAAAAACAACAGCC 5643  
 QY 461 GluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480  
 Db 5644 GAGCTTAATGAAATGACTCAAAAAGCTAACCAAGCTGGGCGCTCAAAATCCTTAAGCGTAT 5703  
 QY 481 AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsnPhe 496  
 Db 5704 GCCGCAATCAGCGGACCATGCTGCGCGAGCATAAACCAATATTTT 5752  
 RESULT 6  
 AP002548/c 325865 bp DNA linear BCT 25-MAY-2002  
 LOCUS Chlamydomphila pneumoniae J138 genomic DNA, complete sequence,  
 DEFINITION section 4/4.  
 ACCESSION AP002548 AB033788 AB033789 AB033790 AB033791 AB033818 AB036083  
 AB036084 AB036085 AB036086 AB036087 AB036088 AB036089 AB038350

AB038351 AB038352 BA000008  
 AP002548.2 GI:10176694  
 Chlamydomphila pneumoniae J138 (strain:J138) DNA.  
 Chlamydomphila pneumoniae J138  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 1  
 Shirai,M., Hirakawa,H., Ouchi,K., Tabuchi,M., Kishi,F., Kimoto,M.,  
 Takeuchi,A., Nishida,J., Shibata,K., Fujiwara,R., Yoneda,H.,  
 Matsushima,H., Tanaka,C., Furukawa,S., Miura,K., Nakazawa,A.,  
 Ishii,K., Shiba,T., Hattori,M., Kuhara,S. and Nakazawa,T.  
 Comparison of outer membrane protein genes omp and pmp in the whole  
 genome sequences of Chlamydia pneumoniae isolates from Japan and  
 the United States  
 J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)  
 20298986  
 2  
 Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Kishi,F., Ouchi,K.,  
 Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.  
 Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 from Japan and GMD029 from USA  
 Nucleic Acids Res. 28 (12), 2311-2314 (2000)  
 20330349  
 3 (bases 1 to 325865)  
 Shirai,M.  
 Direct Submission  
 Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University  
 School of Medicine, Department of Microbiology; 1-1-1  
 Minamikogushi, Ube, Yamaguchi 755-8505, Japan  
 (E-mail:mshirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,  
 Fax:81-836-22-2415)  
 On or before Sep 15, 2000 this sequence version replaced  
 gi:6172302, gi:6172304, gi:6172306, gi:6172308, gi:6172400,  
 gi:6635182, gi:6635184, gi:6635186, gi:6635188, gi:6635190,  
 gi:6635192, gi:6635194, gi:8547442, gi:8547447, gi:8547451,  
 gi:8979170.  
 Location/Qualifiers  
 1..325865  
 /organism="Chlamydomphila pneumoniae J138"  
 /strain="J138"  
 /db\_xref="taxon:138677"  
 /note="synonym:Chlamydia pneumoniae (strain J138)"  
 221..1318  
 /gene="Cp10797"  
 221..1318  
 /gene="Cp10797"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAA99005.1"  
 /db\_xref="GI:8979171"  
 /db\_xref="GI:8979171"  
 /translation="MSKRIKIVLCHLITCLTFRGVLCNAALSNIGVSTSOEFPYOKSI  
 EDMKGYTFDLELLSKGMSERAIISNGSRIVGASGAGSVAYIWESH.LKHLGT  
 LGSEASAREGISKDEGVYVWSDTRREGYKIAFYFDGDMDLDTLAIYSVAVSGVD  
 GSIILVGSATFARQEDGVMGYVKEKRIKQLDLPOGLSEANALISDGTIVLGKE  
 ISRNHIVAYKWNNAVYSLGTLGGSVASAIAISANKVIVGSETNNNGEHNHMHDE  
 TMHDLGTLGGGSPVATGVSADGAIYFSGVSKVGEIHAFYTAGEMHDLTTLGGEAR  
 VPDISEGNDITIGIKTDAERAYLFHHK"  
 1467..2480  
 /gene="Cp10798"  
 1467..2480  
 /gene="Cp10798"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAA99006.1"  
 /db\_xref="GI:8979172"  
 /translation="MKKTCQNRISGIVPSVYLFTVTTQTFRAGHFIDTIGSLYSM  
 ARVSGDGRVYVTEGGNAKRYVDGEFTLEGLVPSERALVFRASDGSYITIGISDD  
 PSORAVKYNVAGALVDLGIIFSEGMQSFAGVSSDGKTIIVGCLYSDDTETNFAVWDEG  
 MVVLPNLPEDRHSICAMDASDEGSVIIVGDAMGSEIARVAVWKGDEHLLINFGAKRS  
 SAHVSXKDSFIYGEIFSENEVHAIFYHNGVYIKDITGLGDSVATGVSRGDKVIVG  
 HSNRTDGEYARFYVVDGRMIDLTIGLSASFARVGSDDGKTIYGRKETELGECNFIY

gene  
CDS  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF38834.1"  
/db\_xref="GI:7189975"  
/translation="MNKKPKRTKRAVOSKAPVKKVPPEESQEAIIQOLTELAVSLPYKE  
LPIAQTFASLIDKNOINSTIALSGTLESLHLELTGCLPFSQEDANFAKELSSVYH  
GLKRLITVYNNQMYKGAE"  
4273..5739  
/gene="CP1062"  
4273..5739  
/gene="CP1062"  
/note="conserved hypothetical protein; identified by  
glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF38835.1"  
/db\_xref="GI:7189976"  
/translation="MSISSSSGPDNKNIMSOVLITSPGGVQDRLSGNETKQIQDT  
RQGTKEDESAITAGAGKDKTSTSTTFETPAPOGVAAKESSESQKAGADVGSGA  
AATYSNATKIAMOTLIEEAKSMESTLEISQSLSAOMKEVAVYVAALSGSGKSG  
ALKLEPRLPKRGVMPRESVITELGALAKAIQTLGAEATRSALSNVAOQAODNRKKG  
LEKAKIKDKREHEQEKKADEKSKDEGMDPVNVTAVSAITVISAIIIPFC  
GAGLGLAGAAVGAAGAAAGAAATVATQTVQAVOVAQAVITAVRQDITIAA  
IKAAVKSIIKAFIKTLVRAIAKAIKSGISKSVFAGTOMIANFRLSKVYSLTSKRW  
TVGVVVAALPDLGKIMOMOLEMOQVAVQFQKEVCKLQAADMTSMFTQFMQOASK  
IASKOTGSSNMPTQKATKLTGAQILKAVAGISATAGAA"  
5772..7106  
/gene="CP1063"  
5772..7106  
/gene="CP1063"  
/note="conserved hypothetical protein; identified by  
glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF38836.1"  
/db\_xref="GI:7189977"  
/translation="MTSGVSGSSSODPTLAAOLAOSQKAGNAOSGHDTKVNTQGAQ  
AEVAGAGGEDILIDQASQGTKEATSTTSKSGEKSEKSGKSTSVASAEATDN  
AAVAGPGLGEDIIDQASQGTKEATSTTSKSGEKSEKSGKSTSVASAEATDN  
KASDSONQAIRSOYESAPATGEAIKROANHOASATEAOASLISGLVNVGTTSV  
GAGTFSAAKGTASALKSAPKFKETGASAAAGAKALTSASSVQOPIMASTAKAATPV  
AASSASATKKAANTLTDMAAASASMAAGSAGKAGGLFGEVLNKPMMSEKSVSGMNV  
VKTQCARVASPAGNALSSMOMOSLMHGLTAVGLISAGOTGEVAAHNRILAGQAEQ  
AEVLEKMSVYGGQAGQAGCOLQBOAMQSFNALTQINADISQVOTTSATFN"  
complement(7177..8094)  
/gene="CP1064"  
complement(7177..8094)  
/gene="CP1064"  
/note="membrane protein, putative; identified by match to  
PPAM protein family HMM PF00892"  
/codon\_start=1  
/transl\_table=11  
/product="membrane protein, putative"  
/protein\_id="AAF73728.1"  
/db\_xref="GI:8163547"  
/translation="MHSMLFLVPLTAFIMSSSPALSKRYMNASAPITATCARMYIAGA  
ITLAAWRRGGVGIISKIFLITVLLATLGTLYTNIFEFTIQSLSSSKTCTFYGLSP  
LMSALSFYTLQEKVETLKRVLGLSLGLVSYICYLPGGQDSDQPTWQJGLPDLIL  
GAASLSTGLWTLRLQIEROSTLSVTAIINAYAMLIAGLMLSHNSHVAVEWPRLPYDIL  
SOFILVTLAVINSICNLVAKLRLKRSYFLSPCNLYMPLYSGFYGMILLGKGVSS  
LGILVAVFMVAQRCILYHNEERQGIYS"  
8288..10195  
/gene="CP1065"  
8288..10195  
/gene="CP1065"  
/note="similar to GB:M36593 SP:P18256 PID:143764  
PID:1565235 GB:AL009126; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="threonyl-tRNA synthetase"

```

/protein_id="AAF38837.1"
/db_xref="GI:7189978"
/translation="MIQVTDOKNVEYLEGSTAAELAKLNKSHOFIGVLINERPD
LSTHNEDGTFLFSEDESEGEKEIFLTHSAHLAKOAVLWLPDAIPTGVIDGFEYD
FANLSTSEDPLELIEDTVKQIYDEKLAISPTTGDKQALAPQNPFKELLREPE
NEELSAISQGEFFEDLCGPHLPSTAHYKAKAVLTSATYKRQDPSRSLSVRIIGTSP
TSKLRALHLEIEBAKRDRHVLGAKLIDLSQDSSFGMPFFHRGHTVDALLRIRK
OLHTAAYKEELLPOLMNRLMEVSGMDVYKAMWYLIQIDEDYALIKPNKCGCMILY
KYTLHSHYKEEPLFLVAEIVGHVHREASGALGIMLRVAFHQDDAHVLFPEQVEETL
NIIDLVSTVGTEFLYEHLSELTSPREKDTIGDSLMELADALBRALVQSGTPEIYAP
GEGAFVYPKIDIHVKDAIORTMOGCTIOLMPLPERELYTEIAQGRKSVPMIHRAL
FGSLERFGLIENFKRFRPLMLSPEDVRIITYADRIIPRAKELEEMKRLGLVYTD
DSSVSYSKRIKNAQNMQVNYITLGDHEINENVLAVTRNDNRVINDVSEFRLNTILE
EKNSLSTALL"
10228..10995
/gene="Cp1066"
10228..10995
/gene="Cp1066"
/notice="similar to GB:D26185 SP:P37522 PID:467391
PID:580906 GB:AL009126; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=1
/product="Para family protein"
/protein_id="AAF38838.1"
/db_xref="GI:7189979"
/translation="MKTIIVNSFKGTAKTSTLHLGAALAOYHOARVLLIDPDAON
LTSGLGDPDCYDGLAVVLQGEKIEQEVIRIPDTOTDLIPADPWLERIEVSGNLAD

```

Alignment Scores:

Pred. No.:	4,74e-118	Length:	15455
Score:	2356.00	Matches:	496
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	0
Query Match:	99.45%	Indels:	1
Db:	1	Gaps:	0

US-09-889-314-2 (1-496) x AE002263 (1-15455)

```

QY      1  ASPTRRASMEISERILSESERSESGERGYPPOASPAINGLYSASNIIEMETSER 20
Db      4264  GATACAAACATGTCATATTCATCTTCAGAGACCTGACAAATCAAAAAATATCATGTGCT 4323
QY      21  GINVAIIeuthrSerThProGInGlyVaProGInGInAsplyLseuSerGlyVaSnGlu 40
Db      4324  CAAAGTTCGACATCGACACCCAGGGCGGTGCCCAACAAGTAACTGTCTGCGCAACGAA 4383
QY      41  ThrLySGInIeGInGInThrArgSInGlyLysAsnThrGluMetGluSerAspAlaThr 60
Db      4384  ACGAAGCAAAATACACCAACACAGCTCAGAGSTAAACACATCGAGATGCAAAAGCATGCCACT 4443
QY      61  IIEAlaGlyAlaSerGlyLysAsplyThrSerSerThrThrLySthGluThrAlaPro 80
Db      4444  ATTGCTGGTGGCTTCGGAAAAACAAAACTTCCTCGCATACAAAAACGAAACACCTTCCA 4503
QY      81  GInGInGlyVaIAlaAlGlyLysGluSerSerGluSerGInLySAlaGlyAlaAspThr 100
Db      4504  CACAGCGGAGTTGCTGCTGGGAAGAAATCTTCAGAAAGTCAAAAGCGAGGTGCTGATCT 4563
QY      101  gLyVaISerGlyAlaAlaAlaThrThraLaseAsnThrAlaThrLySIIeAlaMetGln 120
Db      4564  GGAGATATCAGGAGCGGTGCTACTACACATCAATATCTGCAACAAAAAATTGCTATGCAG 4623
QY      121  ThrSerIIEgluGluAlaSerLySerMetGluSerThrLeuGluSerLeuGInSerLeu 140
Db      4624  ACCCTATTGGAAGGCGAGCAAAAGTATGGAGTGTACTTACCTTAAGATCATCTTCAAGCCTC 4683
QY      141  SerAlaAlaGInMetLySGluVaGInuAlaVaIAlaAlaIeueSerGlyLysSer 160
Db      4684  AGTGGCGGCGCAATATGAAGAACTCCAAGCGGTGTGTGTGCTGCCCTCTCAGGGAAGAAAGT 4743
QY      161  SerGlySerAlaLySLeuGluThrProGluLeuProLySProGlyVaIThrProArgSer 180

```







Db 1247 GCTGTCAATCTGGAAATTAAGACATTTATATCAAACTTTAGTCAAGCGATTGCCAAAGCC 1306  
Qy 361 ILESerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspHe 380  
Db 1307 ATTTCTAAGGAATCTCTAAGGTTTTCCTAAGGAACTCAATGATGCGAAGACCTTC 1366  
Qy 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGly 400  
Db 1367 CCCAAGCTCTCGAAGATCATCTCGCTCTTAOCACATAATGGGTACAGGTTGGGTTGG 1426  
Qy 401 ValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMet 420  
Db 1427 GTTGTAGTTGGCGCCCTCTCGTAAAGGATTATGCAAAATGCACCTCTCGAGATG 1486  
Qy 421 GlnGlnAsnValAlaGlnPheGlnLysGlnLysGlyValGlyLysLeuGlnAlaAlaAspMet 440  
Db 1487 CAACAAACGTCGCTCAATTTAGAAAGAGTCGGAAACATGCAAGCTGCGGCTATATG 1546  
Qy 441 ILESerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460  
Db 1547 ATTTCTATGTTCACTCATTTTGGCAACAGCACTAAATATGCTCAAAACAAACAGGC 1606  
Qy 461 GlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480  
Db 1607 GAGTCTAATGAATGACTCAAAAGCTACCAAGCTGCGGCTCAATCTTAAGCGTAT 1666  
Qy 481 AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAspHe 496  
Db 1667 GCCGCAATCAGCGGAGCCATGCTGCGCAGCATAAACCAATATATTT 1715

RESULT 4  
EL6674  
LOCUS EL6674 1939 bp DNA linear PAT 28-JUL-1999  
DEFINITION ~~GenBank~~ coding Chlamydia pneumoniae antigen polypeptide.  
ACCESSION EL6674  
VERSION EL6674.1 GI:5711357  
KEYWORDS JP 1998212298-A/4.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE  
1 (bases 1 to 1939)  
Izutsu,H. and Hagiwara,T.  
HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION  
PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN,  
RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE  
RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND  
PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY  
Patent: JP 1998212298-A 4 11-AUG-1998;  
HITACHI CHEM CO LTD  
JOURNAL  
OS Chlamydia pneumoniae  
PN JP 1998212298-A/4  
PD 11-AUG-1998  
PI 31-JAN-1997 JP 1997018522  
PI IZUTSU HIROSHI, HAGIWARA TOSHIKATSU  
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC  
C12P21/08,  
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC  
C12N1/91),  
PC (C12P21/02,C12N1/91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH key Location/Qualifiers  
FT source 1. 1939  
FT /strain='YK-41' /organism='Chlamydia pneumoniae' FT  
FT 5'UTR 1..235  
FT CDS 236..1702  
FT /product='chlamydia pneumoniae antigen' FT  
FT polypeptide'

FEATURES FT 3'UTR 1703..1939.  
source 1. 1939  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 631 a 433 c 450 g 425 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,91e-119 Length: 1939  
Score: 2356.00 Matches: 496  
Percent Similarity: 99.80% Conservative: 0  
Best Local Similarity: 99.80% Mismatches: 0  
Query Match: 99.45% Indels: 1  
DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x EL6674 (1-1939)

Qy 1 AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
Db 227 GATCAAAACATGCTATTTCAATCTCTTCAGGACCTGCACATCAAAAATATATCATGTCT 286  
Qy 21 GlnValIleuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40  
Db 287 CAAGTTCTGACATCGACACCCAGGCGTGCCTCCCAACAAAGATTAACCTGCTGGCAACGAA 346  
Qy 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60  
Db 347 ACGAAGCAAAATACACCAACACGTCAGGTTAAACACTGAGATGGAAGCGATCCACT 406  
Qy 61 ILeAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80  
Db 407 ATTGCTGGTGCTCTTGGAAGACAAAACCTCTCTGACATCAAAAACAGAAACAGCTCCA 466  
Qy 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100  
Db 467 CAACAGGAGGAGTGTGCTGGGAAAGATCTTCGAAAGATCAAAAGCAGAGCTGATACT 526  
Qy 101 GlyValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
Db 527 GGAGTATCAGAGCGGCGTCTACTACAGCATCAAAATCTGCAACAAAATTCCTATGCAG 586  
Qy 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140  
Db 587 ACCTCATTTGAAGAGCGAGCAAAAGTATGAGTCTACTTAGAGTCACTTCAAAAGCCTC 646  
Qy 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaAlaAlaLeuSerGlyLysSer 160  
Db 647 AGTGGCGCGCAAAATGAAGAAAGTCAAGCGGTTGTGTGCTGCGCTTCACGGGAAAGT 706  
Qy 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
Db 707 TCGGGTCCGCAAAATGGAAACACCTGAGCTCCCAAGCCCGGGGGTGCACACAAAGATCA 766  
Qy 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr 200  
Db 767 GAGGTTATGCAATATGCAATCGCGCTGCTTAAGCATTCACACATTTGGGAAACAGCACA 826  
Qy 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
Db 827 AAATCTGCTTATCTTAACATGCAAGTACACCAAGCAACAGCAGCAACCAATTAACCTA 886  
Qy 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnLys 240  
Db 887 GGTCTAGAAACCAAGCGGTAAAAATCGATTAAGACGAGGAAGATACCAAGAGATGAAG 946  
Qy 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260  
Db 947 GCTGCGCAACAGATCTAAAGATCTCGAAGAACATGATAGTCAATCTGATGATG 1006  
Qy 261 ILeAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrGly 280  
Db 1007 ATCGGGTTTGTGTCCTACATGATTTCTATTGTGTGCTATTTTCTACATGCCGA 1066

DB 2814 CCCAGCTCTGGAAGTCATCTGCTCTTACCAAGTAATAGGCTCACGGTTGGGTTGGG 2755  
QY 401 ValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMet 420  
DB 2754 GTTGTAGTTGGCGCGCTGCTCGGTAAGGGATTATGCAATTCGAGCTCTCGGAATG 2695  
QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAspMet 440  
DB 2694 CAAACAAACGCTGCTCAATTTCCAGAAAGAGTCGAAACATGCAAGCTGCGCTGATATG 2635  
QY 441 IleSerMetPheThrGlnPheThrPrgGlnAlaSerLysIleAlaSerLysGlnThrGly 460  
DB 2634 ATTTCATGTTCTCACTCAATTTTGGCAACAGGCAAGTAATAATGCTCAAAACAAACAGG 2575  
QY 461 GluSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThr 480  
DB 2574 GAGTCAATGAATGACTCAAAAGCTACCAACACTGGCGCTCAATTCCTTAAGCGTAT 2515  
QY 481 AlaAlaIleSerGlyAlaIleAlaGlyAlaIleHisLysThrAsnAspMet 496  
DB 2514 GCGCAATCAGCGAGCAGCATCGCTGGCGCACATAAAACCAATATTTT 2467  
RESULT 3  
E16639  
LOCUS E16639 1939 bp DNA linear PAT 28-JUL-1999  
DEFINITION Chlamydia pneumoniae gene for 53kDa-antigen peptide.  
ACCESSION E16639  
VERSION E16639.1 GI:5711322  
KEYWORDS JP 1998210978-A/5.  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1939)  
AUTHORS Izutsu, H. and Hagihara, T.  
TITLE RECOMBINANT VECTOR AND TRANSFORMANT CONTAINING THE SAME, AND  
CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE  
PATENT: JP 1998210978-A 5 11-AUG-1998;  
HITACHI CHEM CO LTD  
COMMENT OS Chlamydia pneumoniae  
PN JP 1998210978-A/5  
PD 11-AUG-1998  
PF 31-JAN-1997 JP 1997018523  
PI 1ZUTSU HIROSHI, HAGIMARA TOSHIKATSU  
PC C12N1/09, C07H21/04, C07K14/295, C12N1/21, C12N5/10, C12N7/00, PC  
C12P21/02//  
PC A61K39/00, A61K39/118, C12Q1/68, (C12N1/09, C12R1:01), (C12N1/21,  
PC C12R1:19),  
PC (C12N5/10, C12R1:91), (C12N7/00, C12R1:92), (C12P21/02, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key Location/Qualifiers  
FT source 1. 1939  
FT /strain='YK-41', /organism='Chlamydia pneumoniae' FT  
FT 236..1702 /tissue\_type='elementary bodies' FT CDS  
FT /product='53kDa-antigen peptide'.  
FEATURES  
source Location/Qualifiers  
1..1939  
/organism="unclassified"  
/db\_xref="taxon:32644"  
BASE COUNT 631 a 433 c 450 g 425 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.91e-119 Length: 1939  
Score: 2356.00 Matches: 496  
Percent Similarity: 99.80% Conservative: 0  
Best Local Similarity: 99.80% Mismatches: 0  
Query Match: 99.45% Indels: 1  
DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E16639 (1-1939)  
QY 1 AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
DB 227 GATACCAACATGCTATTTTCATCTTCTTCAGACCTGACAAATCAAAAAATATCATGCTC 286  
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40  
DB 287 CAAGTTCTGCATCGACACCCAGGCGGTCCCCACACATATAGCTGTGGCAACGAA 346  
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60  
DB 347 ACCAAGCAATATACAGCAACACAGTCAGGGTAAACACTGATGGAAGCGATGCCACT 406  
QY 61 AlaAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaPro 80  
DB 407 ATGCTGGTGGCTTCTGGAAAAAGCAAAACCTTCGCTCAAAAAACGAAACAGCTCCA 466  
QY 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThr 100  
DB 467 CAACAGGAGTTCTGCTGGGAAAAAGATCTCTCAAAAGTCAAAAGCGAGTGTATCT 526  
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
DB 527 GAGTATACAGAGCGGCTGCTACTACAGCATCAATATCTGCAACAAAATGCTATGACG 586  
QY 121 ThrSerIleGlnLysAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140  
DB 587 ACCCTATATGAAGAGCGAGCAAAAGATGAGTCTTCACTTCAAGTCACCTCAAAACCTTC 646  
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSerGlyLysSer 160  
DB 647 AGTGGCGGCAATGAAAGAGTCGAAGCGGTGTTGTTGCTGCTTCAGGAAAAAGT 706  
QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProGlyValThrProArgSer 180  
DB 707 TCGGTTCCGCAAAATTTGGAACACCTGAGCTCCCAAGCGGGGTGACACCAATGCA 766  
QY 181 GluValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyAlaThr 200  
DB 767 GAGTTATGCAAAATCGAGCTCGGCTTCTTAAAGCAATTCAGACATGGAGAGGCACA 826  
QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
DB 827 AATCTGCTTATCTACTATGTAATACACAGCAACGACCAACCAAAATTAACATCA 886  
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysMetLys 240  
DB 887 GGTCTGAAAGAGCAAGCATTAATAATCGATTAAGAACGAGAAATACCAAGATGAG 946  
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260  
DB 947 GCTGCGAAGCAGAGCTTAAGATCTCGAAGCAACATGATGATCTGCTGATG 1006  
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280  
DB 1007 ATCGCGGTTCTGTTCCCTTACAGTATTTCTATTGCTGCTGCTATTTTTCATGCGGA 1066  
QY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaAlaAlaAlaAlaGlyAla 300  
DB 1067 GCTGAGCTCGGTGAGTCTGCTGGGAGCTGCTGAGTGAAGCGGCGACCTGGAGTGA 1126  
QY 301 AlaGlyAlaAlaAlaAlaAlaThrValAlaAlaThrGlnIleThrValGlnAlaValGln 320  
DB 1127 GCAGGAGCTGTGCGCAACACAGGTAGCAACACAAATTAAGTTCAAGCTGTGCCAA 1186  
QY 321 AlaValLysGlnAlaValIleThrValAlaArgGlnAlaIleThrAlaAlaIleLysAla 340  
DB 1187 GCGGTGAACAAAGCTGTATTCACAGCTGCACAAACAGCATCCGCGCTTAAGAGG 1246  
QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360

```
CDS
8267..9520
/feature="Cp0814"
/codon_start=1
/transl_table=11
/product="CT573 hypothetical protein"
/protein_id="AAD18952.1"
/db_xref="GI:4377126"
/translation="MAGSPYVVDLYRISLSAESLGGIRAPQAESLLGCALANF
PDLEERLPDLRKLLEFLGSDNDPACGKGFQJLASKECYIALKREVLNVTSSR
GPAVSESPKGVPELIMIECSVSDGSRVAVRLOGJLHKLISPRDDETLFINRAN
KLDCWEINAGFRVDASFPYKOKIRIRGVKPLIMHGAGATKATKTERDVPSSDENY
SRVLAVGVLLMDGNCWQTCGFQGASSRAPLFEVKRIDDKVMTADLNNVGTOQKTL
SLVKGVPSPRIENIEVRIEFTGMRSMKPTLVGGORLILSPDMVILRTAKGMKLS
RADIOIDYVTKGVTKGPIILVFPEKLEKDLRGFVLRGHMFNAORTIVETISLPKQGFEP
VASQESVNTSRSSAAAHPGATNRGGS"
9521..11785
/feature="gspe"
/feature="Cp0815"
/feature="gspe"
9521..11785
/feature="gspe"
/feature="gspe"
9521..11785
/transl_table=11
/codon_start=1
/product="General Secretion Protein D"
/protein_id="AAD18953.1"
/db_xref="GI:4377127"
/translation="MVFPRNSLHLHVALSGMLCCSSGVALRTIAEKMASLHSGRGAD
YEGASFANMREYSIOISKLYEARKIYSEARKEALMDKILIRIGVRYLRIEE
LMAAIEIRKGNLEDEYALMNHPEITIVLVVDYGDDESIYILPOIGIKRATLSKV
VPESEFDECLQILSRIGIVRQVNSMKEILMKKECSVAAGVSSKRLDEALPETA
YIGFVLNSVDVAHNLQHLKFIINETHVDVIAKRWI FSGAGVSKDLITYNVQS
ESIRQEVVILPLTKIDPCGEMISILNAAFREVDITKQVSEESGLRVVPLOYGRSLFS
GTAAIVQALITLIRELEGIEINPTKTVFVYVKHSDPOELALLSOVHDFVSGENKA
SVGADGCGSOLNASIOIDITVYSSAKGSAVYKGNFIDSKGTILIMVEXEVLPTIO
MLKRLDYPKMWRIEVLIFPERKILABQSGINLRIGECYOKKCGSSPSMAGCTG
LELFKSGTSSIVPGDYLAOFLMAQEDVRINASPVTYVNMOTPARIAVDEMIAV
SSDKKAQYNRAQVGIKIMKLPVINVGEDKSYITLTEDITFGNHDPRPVT
RNITNKVRIADGEVITIGLRCKWMSDSDHGIPLGDIPLGIGKLFMGSTSDSLEMF
VEITPKILENVEODEERKEEALLSRPEREYVQALAASEAAARAHHKLEMPASG
VLSOVEROEYDGC"
11775..13265
/feature="gspe"
/feature="Cp0816"
/feature="gspe"
11775..13265
/feature="gspe"
/feature="gspe"
11775..13265
/transl_table=11
/codon_start=1
/product="General Secretion Protein E"
/protein_id="AAD18954.1"
/db_xref="GI:4377128"
/translation="MAASILSELELDILPYFLKHKCLPIEESSEATIIAHATATSV
IAODEVKLLIKRPVRFVLEKESEILQRLQOLVSNREGVNSDKLTMKREDDGTISEE
DLLETEDTIPVYRLNLWLKKAIEERADIHPEPEDSMRLRYRIDVULHHRHSPSH
LRSALTRKVLAKMDIAEHLRPQDGRKIRIHGGQEVDMKSTVYITIGERVALLD
KRNVLIDLAGLHMPKGETILFEKDTTAAEGILLVPGSGEATTLVLEKLPILT
NIMTIEDPEPKLPGIAQIAVAPKIGLTFARGLRHLRQDPIILVWGEI.RDQETAEIA
IOAALTGLHVVSTLHTNDALISAIPRLLMGIESYLSLTIIVVVAQRLVRIETCPCKV
ATPENQKSFSLASGKDPTEMLPYRGGCVHCFRSGYGRGQIYEFLPNLIFRBEVA
```

## Alignment Scores:

Pred. No.:	1,05e-118	Length:	16903
Score:	2369.00	Matches:	496
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-889-314-2 (1-496) x AB001662 (1-16903)

QY 1 AspThrAsnMetSerIleSerSerSerGlyProAsnGlnLysAsnIleMetSer 20  
DB 3954 GATACAAACATGCTATTCCTTCAGGACCTGACACATCAAAAATATCATGCTCT 3895  
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40

```
|||||
Db 3894 CAAGTTCTGACATCGACACACCCAGGCGCGCCCAACAAGATAAAGCTGCTGGCAACGAA 3835
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyAsnThrGluMetGluSerSpsalThr 60
Db 3834 ACGAAGCAAAATACAGCAAAACAGCTCAGGGTAAACACATGAGATGGAAAGCATCCACT 3775
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80
Db 3774 ATTGCTGGTCTTCTTGGAAGAACAAAACCTTCTGACATCAAAAAACAGAACAGCTCCA 3715
QY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
Db 3714 CAACAGGAGGATTGCTGCTGGAAAGAAATCCTCAGAAAGCAAAAGGACAGCTGCTACT 3655
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
Db 3654 GGAGTATCAGAGAGCGGCTCTACTACAGCATCAAAATACGACAAAATTCCTATGACAG 3595
QY 121 ThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu 140
Db 3594 ACCTCATTTGAAGAGCGAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCTTC 3535
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSer 160
Db 3534 AGTGGCGCCCAATGAAAGCAAGTCGAAGCGGTGTTGTTGCTGCTGCTCAGAGAAAGCT 3475
QY 161 SerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProAsnSer 180
Db 3474 TCGGGTTCGCAAAATATGGAAACACCTGAGCTCCCAACCCGCGGGTGACCCAAACATCA 3415
QY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr 200
Db 3414 GAGGTATATGAAATCGGACTCCGCTGCTTAAAGCAATTCAGACATTTGGAGAGACGACA 3355
QY 201 LysSerAlaLeuSerAsnTyraIaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 3354 AATCTGCTTATCTAATCATGCAAGTACACAAGCAAGCAAGCAACCAATAACTCA 3295
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlnTyrcGlnLysMetLys 240
Db 3294 GGTCTGAAAGCAACGACGATTAATAAATCGATTAAGAAAGCAAGAAATCAAGAGATGAAG 3235
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db 3234 GCTGCCGAACAGAGCTTAAAGTCTCGAAGGACACATGATCTCATATCTGATGATG 3175
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIleIleThrCysGly 280
Db 3174 ATCGCGGTTTCTGTTCCATTCACAGTTATTTCTATTGTTGCTGCTATTTTACATCGCGA 3115
QY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAla 300
Db 3114 GCTGGACTCGCTGGAGTCTCGCTGGGAGCTCGTGAAGGCGACCGGCTGAGGTGA 3055
QY 301 AlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln 320
Db 3054 GCAGAGACTGCTGCCCAACCAAGGACGACGACCAAAATTTACACTTAAGCTGTGTCGCA 2995
QY 321 AlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAla 340
Db 2994 GCGGTGAACCAAGCTGTATTCACAGCTGTACAGCAAGCATCACCGCGCTATTAAGAAGCG 2935
QY 341 AlaValLysSerGlyLysAlaIleLysAlaIleLysThrLeuValLysAlaIleAlaLysAla 360
Db 2934 GCTGTCAAAATCTGGAATTAAGCATTTATTCAAAACCTTATGTCAAAACGATGGCAAAAGCC 2875
QY 361 IleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
Db 2874 ATTTCTAAAGGAATCTTAAGGTTTTCGCTAAGGAACTCAAAATGATTGCGAAGAACTTC 2815
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTryValThrValGlyValGly 400
|||||
```



AX033283  
LOCUS AX033283 1491 bp DNA linear PAT 21-SEP-2000  
DEFINITION Sequence 1 from Patent WO0046359.  
ACCESSION AX033283  
VERSION AX033283.1 GI:10280098  
KEYWORDS  
SOURCE Chlamydomophila pneumoniae.  
ORGANISM Chlamydomophila pneumoniae.  
REFERENCE Bacteria: Chlamydiales; Chlamydiaceae; Chlamydomophila.  
AUTHORS Matthews, R.C. and Burnie, J.P.  
TITLE 1 (bases 1 to 1491).  
JOURNAL Patent: WO 0046359-A 1 10-AUG-2000;  
MATTHEWS RUTH CHRISTINE (GB) ; NEUTECH PHARMA PLC (GB) ; BURNIE JAMES P (GB)  
FEATURES  
source location/Qualifiers  
1..1491  
/organism="Chlamydomophila pneumoniae"  
/db\_xref="taxon:83558"  
<1..1491  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC09981.1"  
/db\_xref="GI:10280098"  
/translation="DTNMSISSSGPDNOKNIMSOVLSTPOGVPOODKLSGNETKOI  
OOTROGKNTMESDATTAGASGDKXTSTTETETAPOGVAAKESSESOKAGADTV  
SGAAATTSRTATRTIAMOTSTIEBASKSHESTLESIOSSAOMKEVEAVVAALSGKS  
SGSAKLETPELPKGVTPRSREIVEIGLALAIQTLGATTSALSNASTOQADQTN  
KLGLEKQAIKIDKEREVEYEMKAEQKSKDELTGNTVAVVAIVSAIVTIVSAI  
FTCGAGLAGLAGAAGVGAAGAAGAAATVATQIVQAVOVAOKAVTAVFOAI  
TAAIKAAVKSIGKAIKTLVKAIAKAIKSKISKSTOMIAKPELSKIVISLTS  
KWTIVGVGVVVAAPALGKIGIMOMOLEMOQNVADQKREVGTLQAAADISMFTQWQO  
ASKIASQKSGSNMPOKATKLGAIILKAVAIISALINAKHTNFE"  
BASE COUNT 485 a 335 c 348 g 323 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,64e-120 Length: 1491  
Score: 2369.00 Matches: 496  
Percent Similarity: 100.00% Conserved: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-889-314-2 (1-496) x AX033283 (1-1491)  
QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
DB 1 GATACAAACATGCTATATTCATCTCTTCAGGACCTGACATCAAAAATAATCATGTCT 60  
QY 21 GlnValIleuThrSerThrProGlnGlnValProGlnGlnAspLysSerGlnGln 40  
DB 61 CAAAGTTCGACATCGACACCCAGGCGGCGCCCAACAAAGATAGCTGCTGGCAACGAA 120  
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThr 60  
DB 121 ACGAAGCAAAATACAGCAACACGCTGAGCTGAGCAATGCAATGCAATGCGATGCTCACT 180  
QY 61 IleAlaGlnLysSerGlnLysAspLysThrSerSerThrThrLysThrGlnLysAlaPro 80  
DB 181 ATTGCTGGGTCTTGGAAGAACAAACTTCCTCGACTACAAAACAGAACAGCTGCA 240  
QY 81 GlnGlnGlnValAlaAlaGlnLysGlnSerGlnSerGlnLysAlaGlnLysAlaAspThr 100  
DB 241 CAACAGGGAGGTTGGTGGGAAAGATCTCGAAGAACAAAGGACAGCTGCTGATCT 300  
QY 101 GlnLysSerGlnLysAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
DB 301 GGAGTATCAGAGGCGGCTCTCTACAGCATCAATACGCAACAAATTTGCTATGCG 360  
QY 121 ThrSerIleGlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140

DB 361 ACCCTATTTGAAGAGCGACGCAAAAGTATGAGCTACCTTAGACTTCACCTTCAAGACCTC 420  
QY 141 SerAlaIleGlnMetLysGlnValGlnValAlaValAlaIleAlaIleuSerGlnLysSer 160  
DB 421 AAGTCCGCGCAATGATAAAGATGCAAGCGGTTGTTGCTGCGCTTCACAGGAAAT 480  
QY 161 SerGlnSerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSer 180  
DB 481 TCGGCTCCGCAAAATTTGAAACACCTGAGCTCCCAAGACCGGGGTGACACCAAGATCA 540  
QY 181 GlnValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysValThr 200  
DB 541 GAGTTATGCAATGAGATCGCTGCTGCTTAAGCAATTCAGACATTTGGAGAGACGACACA 600  
QY 201 LysSerAlaLysSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
DB 601 AATCTGCTTATCTTACATCTCAAGTACACAGACCAAGACCAACAAATTAACCA 660  
QY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysLeu 240  
DB 661 GGCTTAGAAACCAAGCGCTAAATTCATTAAGAACGAAAGCAATACCAAGATGAAG 720  
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlnLysMetAspThrValAsnThrValMet 260  
DB 721 GCTGCCAAGACAGATCTTAAGATCTCGAAGAACCAATGATGATCAATGCTGATG 780  
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrGly 280  
DB 781 AATCGGGTTCGTTGCCATTCATCTTATTTGTTGCTGCTATTTTTCATGCGGA 840  
QY 281 AlaGlnLysLeuAlaGlnLysLeuAlaGlnLysAlaValAlaGlnAlaAlaIleGlnLysAla 300  
DB 841 GCTGAGCTGCTGAGCTGCTGCTGAGAGCTGCTGAGAGCTGAGAGCTGAGAGCTG 900  
QY 301 AlaGlnAlaAlaAlaIleThrThrValAlaIleThrGlnIleThrValGlnAlaValGln 320  
DB 901 GCAGAGAGCTGCTGCCGCAACACGAGTAGCAACAAATTAACAGCTGCTGCTGCCAA 960  
QY 321 AlaValLysGlnAlaValIleThrThrAlaValAlaGlnAlaIleThrAlaAlaIleLysAla 340  
DB 961 GCGGTGAACCAAGCTGTTATTCACAGCTGTCAGACCAAGCATACCGCGCTTAAAGGG 1020  
QY 341 AlaValLysSerGlnLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
DB 1021 GCTGTCAAATCTGGAATTAACATTTATCAAACTTTATGCTCAAGCGCTGCCAAAGCC 1080  
QY 361 IleSerLysGlnLysSerLysValPheAlaLysGlnLysGlnMetIleAlaLysAsnPhe 380  
DB 1081 ATTCTTAAAGGAATCTCTTAAGGTTTCCGCTAAGGCAACTCAATGATTCGAAAGAACTTC 1140  
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlnValGln 400  
DB 1141 CCCAAGCTCTCGAAGATCTCTGCTCTTACCAGATTAATGGGCTACAGGTGGGTGGG 1200  
QY 401 ValValAlaAlaAlaProAlaLeuGlnLysGlnLysGlnIleMetGlnMetGlnSerGlnMet 420  
DB 1201 GTTGTAGTTGGCGCCCTCTCTCGTAAAGGATTTATCAATGACAGCTTCGCGAGATG 1260  
QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlnValGlnLysLeuGlnAlaAlaAlaAspMet 440  
DB 1261 CAACAAAAGCTGCTCAATTTGCAAGAAAGTCGGAACAACTCAGAGCTGCGCTGATAG 1320  
QY 441 IleSerMetPheThrGlnPheThrPheGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460  
DB 1321 ATTCTTATGTTCACTCAATTTGCAAGAAAGTCGGAACAACTCAGAGCTGCGCTGATAG 1380  
QY 461 GluSerAsnGlnMetThrGlnLysAlaThrLysLeuGlnValGlnIleLeuLysAlaThr 480  
DB 1381 GAGTCTAATGAATATGACTCAAAAAGCTACCAAGCTGGGCGCTCAATCTTTAAACGCTAT 1440  
QY 481 AlaAlaIleSerGlnAlaIleAlaGlnLysLysThrAsnAsnPhe 496  
DB 1441 GCCCAATCAGCGAGCGCATCGCTGCGCACATTAACCAATTAATTTT 1488

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 27, 2003, 14:06:32 ; Search time 2670 Seconds

(without alignments)  
5406.364 Million cell updates/sec

Title: US-09-889-314-2  
Perfect score: 2369  
Sequence: 1 DTNMSISSSGPDNOKNIMS.....LKAYAAISGAIAGAKHTNNE 496

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+pn.model -DEV-rlh  
-Q/cgn2.1/USPTO.spool/US09889314/runat.24012003.144540.10107/app.query.fasta.1.647  
-DB-GenEmbl -QFMT-fastlap -SUFFIX-rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45  
-DOCAIGN=200 -THR.SCORE-pct -THR.MAX=100 -THR.MIN=0 -ALIGN=45 -MODE-LOCAL  
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-US09889314.eccn.1.1.1616/runat.24012003.144540.10107 -NCPU=6 -ICPU=3  
-NO\_XLUPXY -NO\_MMMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pin.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em\_sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2369	100.0	1491	6	AX033283	AX033283 Sequence
2	2369	100.0	16903	1	AE001662	AE001662 Chlamydia
3	2356	99.5	1939	6	E16639	E16639 Chlamydia p
4	2356	99.5	1939	6	E16674	E16674 gDNA encodi
5	2356	99.5	15455	1	AE002263	AE002263 Chlamydo
6	2340	98.8	325865	1	AP002548	AP002548 Chlamydo
7	2326	98.2	9595	6	E16678	E16678 All sequenc
8	2320	97.9	11099	6	E16636	E16636 Expression
9	2317	97.8	1587	6	E16677	E16677 All sequenc
10	2316	97.8	1464	6	AR122105	AR122105 Sequence
11	2316	97.8	1464	6	E12520	E12520 DNA encodin
12	2316	97.8	1464	6	E12535	E12535 Chlamydia p
13	2316	97.8	1464	6	E12541	E12541 DNA encodin
14	2316	97.8	1464	6	E16635	E16635 Chlamydia p
15	2316	97.8	1464	6	AR122114	AR122114 Sequence
16	2316	97.8	1947	6	E12525	E12525 DNA encodin
17	2316	97.8	1947	6	E12538	E12538 DNA encodin
18	2316	97.8	1947	6	AR122110	AR122110 Sequence
19	1282	54.1	5658	6	E12523	E12523 Nucleotide
20	1282	54.1	5658	6	E12544	E12544 Nucleotide
21	1282	54.1	5658	6	AR122122	AR122122 Sequence
22	1265	53.4	5438	6	E12527	E12527 Nucleotide
23	1265	53.4	5438	6	E12540	E12540 Nucleotide
24	1265	53.4	5438	6	AR122109	AR122109 Sequence
25	1262	53.3	1048	6	E11479	E11479 DNA encodin
26	1262	53.3	1048	6	E12265	E12265 DNA encodin
27	1262	53.3	1048	6	AR122107	AR122107 Sequence
28	1245	52.6	777	6	E12522	E12522 DNA encodin
29	1245	52.6	777	6	E12543	E12543 DNA encodin
30	1245	52.6	813	6	AR122106	AR122106 Sequence
31	1245	52.6	813	6	E12521	E12521 DNA encodin
32	1245	52.6	813	6	E12542	E12542 DNA encodin
33	1245	52.6	1286	6	AR122115	AR122115 Sequence
34	1245	52.6	1286	6	E12526	E12526 DNA encodin
35	1245	52.6	1286	6	E12539	E12539 DNA encodin
36	1245	52.6	1286	6	AE001328	AE001328 Chlamydia
37	1233.5	51.1	10653	1	AE002353	AE002353 Chlamydia
38	1223.5	51.6	9983	1	AC024253	AC024253 Homo sapi
39	238.5	10.1	179553	2	STSPGNS	X82670 S. typhi sp
40	220.5	9.3	5393	6	AR082442	AR082442 Sequence
41	220.5	9.3	5393	6	AE062727	AE062727 Salmonella
42	220.5	9.2	274050	1	SDU66877	SDU66877 Salmonella
43	218.5	9.2	2995	1	STU25631	STU25631 Salmonella
44	218.5	9.2	3622	1	AR082441	AR082441 Sequence
45	218.5	9.2	3622	6	AR082441	AR082441 Sequence

RESULT 1

## ALIGNMENTS

Query Match: 6.968 Indels: 169  
DB: 22 Gaps: 25  
US-09-889-314-2 (1-496) x AAH78245 (1-3725)  
QY 9 SerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleuThrSerThrPro--- 27  
DB 784 ACGGAGGGGGCCCAAGACCG-----GGTAAAGATGATCAACCCCAAC 831  
QY 28 -----GlnGlyValProGlnGlnAsp---LysLeuSerGly-----Asn 39  
DB 832 CAGGGGAAAAAGGTGAGGGGGCCCAACCCAGGGCGAAGAGCGGGGCTCAGAAC 891  
QY 40 GlnThrLysGlnIleGlnGlnThrArg---GlnGlyLysAsnThrGlnMetGlnSerAsp 58  
DB 892 CAGGCCAAAAGGTGAGGGGGCCCAAGACCCAGGGCGAAGAGCGGAG-----GGG 942  
QY 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThr 78  
DB 943 GCCCAGAAATCAGGGGCAAAAGGAGGGGGCCCAAGAACCCAGGGCGAAGAGCGGAGA 1002  
QY 79 AlaProGlnGlnGlyVal----- 84  
DB 1003 GCCCAGAAACCCAGGGCCAAAAGAGAGGAGGCCCAAGATCAGGGTAAAGAGCAGAGGG 1062  
QY 85 AlaAlaGlyLysGlnSerSerGlnSerGln-----LysAlaGlyAlaAspThrGly 101  
DB 1063 GCTCAGGGGCAAAAGGCAAAAGAGATCCCAACCAAGCCAAAAGAGAGGAGCTCCC 1122  
QY 102 ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLys-----IleAla 118  
DB 1123 ATCCAGGGGCAAAAGGAGATTCGGTTCCTAATCAGGGCGACAAAGGTAGAGGTATTACA 1182  
QY 119 MetGlnThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGln 138  
DB 1183 AACCCAGGGGAAAAAGCAGAGGGTCCCGCTGAGAGCCAAAAGAGCAGAGAGG----- 1236  
QY 139 SerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaIleAlaLeuSerGly 158  
DB 1237 -----TCCCCCAACCCAGGCAAAAGGCAAGCGCA-----GCTCCCAATCAGGGT 1281  
QY 159 LysSerSerGlySerAlaLysLeu-----GlnThrProGln 170  
DB 1282 AAAAAGACAGATGCTCTCTCTCTCCAGGGCAGAAATACAGATGTGCCCAGAGCCAGAG 1341  
QY 171 LeuProLys-ProGlyValThrProArgSerGlnValIleGln-----Ile 185  
DB 1342 GCACCAAAACCAAGAGGCTCTCCCAAGAGAAAGTGGTTCAAGAAAAAGGTGAGCT 1401  
QY 185 eGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaIleThrLysSer---AlaIle 204  
DB 1402 GGGCCCCCAGATGCCAGCGCCCTCTCTACCTCCCTACCAAGACGCTGCTCCACGGTT 1461  
QY 204 uSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGly 224  
DB 1462 GGGACCATGTGTCTCAAGAGGGCGAG--GCCAGCGGCTCATCAGATCTCTGCCAGAA 1520  
QY 224 sGlnAlaIleLysIleAsp----- 230  
DB 1521 GGCTGGTGCATTCAAGGACCTGGCACAAGGCCACTCAGAAAGGTGACCTGTGGCGAT 1580  
QY 231 -----LysGlnAlaArgGlnGlnGlnLysAlaAlaGlnLysLysSerLysAs 248  
DB 1581 TCTGAAGCCCGCAGCTGGAAGAGAGAAAAACTGCTGGCCACAGAAACAGAAAGATCGGC 1640  
QY 248 pLeu---GlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIle 267  
DB 1641 TGTCCCAAGAGCAAACTGAGGAGCTCAACAAGAGATG----- 1680  
QY 267 eThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAl 287  
DB 1680 ----- 1680

QY 287 aaIaGlyAlaAlaValGlyValAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAla 307  
DB 1681 -----GCAGCAGAAAAGGCCAAAGACACACCGCGGAGGCC----- 1716  
QY 307 rThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIle 327  
DB 1717 -----AAAGTAAAGACAGCTGGT-----GCCCGGAGCAGGAGAT 1754  
QY 327 eThrAlaValaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys-----SerG 345  
DB 1755 CAGGCTGTGGAGGAGCAGCAGCTCAGCCAGCTACCGGGAGCAGGTGAAGAGGTGACAGA 1814  
QY 345 yLleLysAlaPheIleLysThrLeuValLysAlaIleAlaLys----- 359  
DB 1815 GCTCAGAGGCAAGATTCGAGCTTTCAGAGGACGCTGGAGATGGCCCCAACACGAGCT 1874  
QY 360 -----AlaIleSerLysGlyLysSerLys 367  
DB 1875 GGCCCGCTGCAGAGAGAGAACTCCATCCTGGCGGATGGCTGAACACAGCCAGAGCCA 1934  
QY 367 sValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIle 387  
DB 1935 GGTGAGAGCAAGCAGAACGACAGCTGGCCAAAGCTTCGGCAGAGCTTCAGAGAGTCAG 1994  
QY 387 eSerLeuThrSerLysThrPvalThrVal-----GlyValGlyVal 401  
DB 1995 CAAAGAGCTGTGGAGAGACTCAGAGGCTGTGGCGCAAGTGAAGCAGCAGCGAAGACTT 2054  
QY 401 lValValAlaAlaProAlaLeuGlyLysGlyLleMetGlnMetGln----- 416  
DB 2055 GGAAGCCAAAGCAGAGCTGCTTCGAGAGCAGAGTCTTCGACGTGACAGGCGTCCACAGGA 2114  
QY 417 -----LeuSerGlnMetGlnGlnAsnValAlaGlnPheG 428  
DB 2115 GAGTGAAGAGGCTTCGACAAAGCCTGGACGAGGTCAAGCGGAGGCTGTGCACAGCA 2174  
QY 428 nLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTr 448  
DB 2175 GGGCAGCCAGCCAGCTCGCGGGGATGCCAGAAAGGCCAG-----GAGCAGACA 2225  
QY 448 pGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSerAsnGlnMetThrGlnLys 468  
DB 2226 GCAGCAGATGGCCGAGCTCAGCAGCAAGTTACAGTCTCCGAGGAGAGAGTCCGACAG 2285  
QY 468 sAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAlaIleSerGlyAlaIleAl 488  
DB 2286 ATGCGAGGAGCTG-----AGTGGCTTCACGAGGACGCTCCA 2321  
QY 488 aGlyAlaHisLysThrAsnAsn 495  
DB 2322 GAGGCGCAGGGCGGAGAACTCC 2343

Search completed: January 27, 2003, 15:08:47  
Job time : 3371 secs







```

Db 18606 AAACAGATATTAAGATTAGAGTGAGAAC-----ATGAGTGATGATCAGTA 18659
      |||          :|||:|||||
Qy 61 lleaaglyalaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80
      |||          ||| ||| ||| |||
Db 18660 GTAAATGAGATTAGTTGACGATTAATAAAGCAGAT-----AAACACTGGATCGTTT 18713
      |||          ||| ||| ||| |||
Qy 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyLysAlaSer 100
      |||          ||| ||| ||| |||
Db 18714 GACACAGATTTAGCG-----AAAGCAGGACCAAT--- 18743
      |||          ||| ||| ||| |||
Qy 101 GlyAlaSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIleAlaMet 119
      |||          ||| ||| ||| |||
Db 18744 -----GCAGGCGCGCATTAATAGCGTATAGAGACGCGTATTCGATTTGCTGAGT 18797
      |||          ||| ||| ||| |||
Qy 120 GlnThr-----SerIleGluIleAlaSerLysSerMetGluSerThrIleu 134
      |||          ||| ||| ||| |||
Db 18798 CAAATCGAAGCATTAAGACACGTTTGTAAATGCGTTTAAAGCATGGAGAGCTGCGC 18857
      |||          ||| ||| ||| |||
Qy 135 GluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaValAla 154
      |||          ||| ||| ||| |||
Db 18838 TCAAATGCTTTAAAGCTTAAGCTTAATGCGTGAATGCTTCAATGATACAGCC 18917
      |||          ||| ||| ||| |||
Qy 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysPro 174
      |||          ||| ||| ||| |||
Db 18918 GCTATATCTAAACTGATCA----- 18938
      |||          ||| ||| ||| |||
Qy 175 GlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194
      |||          ||| ||| ||| |||
Db 18939 -----ACAGTAAACCTGGGTTGCTTAAACGCTGCTAAACGCTATATACACGCTAAAG 18992
      |||          ||| ||| ||| |||
Qy 195 ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyAlaSer-----ThrGln 211
      |||          ||| ||| ||| |||
Db 18993 GAACCTGGAAACAAGTATCAAAACACACGCGTTAATATTAAACGCGCTTCTTCAATT 19052
      |||          ||| ||| ||| |||
Qy 212 AlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLys 231
      |||          ||| ||| ||| |||
Db 19053 GCTAAGACAGCTCAAAAGTATGTTGTGTCAGCTTAAAGTATCATATATGTCATTA 19112
      |||          ||| ||| ||| |||
Qy 232 GluArgGluGluThrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluGly 251
      |||          ||| ||| ||| |||
Db 19113 TCCATCCCGCGCAATTAAGCGCTGGATCATTAATTAATCAAGCA---TTAGTAAGT 19169
      |||          ||| ||| ||| |||
Qy 252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271
      |||          ||| ||| ||| |||
Db 19170 AGTTTGCACAGCTAAATCGGCT-----GCTATTCTTTGCTCAACTGTA 19223
      |||          ||| ||| ||| |||
Qy 272 lIeValAlaAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAla 291
      |||          ||| ||| ||| |||
Db 19224 GTTATTTAAAGTATT-----CCAGGAGCTGCT 19250
      |||          ||| ||| ||| |||
Qy 292 ValGlyAlaAlaAlaAla-----GlyGlyAlaAlaGlyAlaAlaAlaAlaThrThrVal 309
      |||          ||| ||| ||| |||
Db 19251 AAACACGCGCGCTACAGCACTGTAACAGTTCTGTAGTACAAACGCGTGTAGTAT 19310
      |||          ||| ||| ||| |||
Qy 310 AlaThrGlnIleThrValGlnAlaValAlaValAlaValLysGln----- 324
      |||          ||| ||| ||| |||
Db 19311 GCTGCTTATATGAGCGTTTAAGAGAACTATTAGCGCTGTAAAGCTATTCTTACGCTTACA 19370
      |||          ||| ||| ||| |||
Qy 325 -----AlaValIleThrAlaValAlaArgGlnAlaAlaIleThrAlaAlaIleLys 339
      |||          ||| ||| ||| |||
Db 19371 AATATCAGCGCATTTAGCATTAAGTAGCGCAATGAATAACAGCTTTTATGCGCTGATACAGC 19430
      |||          ||| ||| ||| |||
Qy 340 AlaAlaValLysSerGly-----IleLysAlaPheIleLysThr-----LeuVal 354
      |||          ||| ||| ||| |||
Db 19431 GCGGCGCAAAACACAGCAACAGTGAATAGCATTTAAACAGCGCTTATGCGCTGTG 19490
      |||          ||| ||| ||| |||
Qy 355 LysAlaIleAlaLysAlaLysSerLys---GlyIleSerLysValPheAlaLysGlyThr 373
      |||          ||| ||| ||| |||
Db 19491 AATATCGGAGCTAAAGCGGACGCAAGCTGTGATTTCAGCATTA-----AAAGGCGCTA 19544
      |||          ||| ||| ||| |||
Qy 374 GlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLys 393
      |||          ||| ||| ||| |||

```

```

Db 19545 GGAACATTCGCAAAACACGCTGTTCTTAATTAA-----AGTGATTAGTAACGGA 19598
Qy 394 TrpValThrValGlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleMet 413
      |||          ||| ||| ||| |||
Db 19599 TTT-----AACGCGCAAAACGCGCGCAAAAGTGACAGGCGCTGGAATGCGT 19646
      |||          ||| ||| ||| |||
Qy 414 GlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLys 433
      |||          ||| ||| ||| |||
Db 19647 GACGACCTTAAATTCAGTTTAAATAAGCCCGCCGCAACAAAGCTCGCTTATGATTCTCAGA 19706
      |||          ||| ||| ||| |||
Qy 434 LeuGlnAlaAlaAlaAspMetIleSerMetPhe----- 444
      |||          ||| ||| ||| |||
Db 19707 TTACGACGACGCTGCGATTATTCAGCAACTAAATAATGTTGCGTAGCCGCTATGCT 19766
      |||          ||| ||| ||| |||
Qy 445 ---ThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGluSerAsn 463
      |||          ||| ||| ||| |||
Db 19767 CGAGTTGATACGATTGATACCTAAATGCTTAAACGCTTACAGCTTCCTGCTTACACAAA 19826
      |||          ||| ||| ||| |||
Qy 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaLysAlaAlaIle 483
      |||          ||| ||| ||| |||
Db 19827 GATGCGCAACTGATTATGACACACTTACAGCCGCTATGATGCGACACCAATCCGATTA 19886
      |||          ||| ||| ||| |||
Qy 484 SerGlyAlaIleAlaGlyAlaHisLys 492
      |||          ||| ||| ||| |||
Db 19887 GATGCTGTGCAATTAGTGCTTAAATA 19913
      |||          ||| ||| ||| |||

RESULT 44
AB069245
ID AB069245 standard; DNA; 3011208 bp.
XX
XX
AC AB069245;
XX
XX
DT 29-AUG-2002 (first entry)
DE
DE Listeria innocua DNA sequence #684.
XX
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria innocua.
OS
XX
XX WO200228891-A2.
XX
XX PD 11-APR-2002.
XX
XX PF 04-OCT-2001; 2001WO-FR03061.
XX
XX PR 04-OCT-2000; 2000FR-0012697.
XX
XX PA (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX PI Kunst F, Glaeser P;
XX
XX DR WPI: 2002-332479/37.
XX
XX PT New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX
XX PS Claim 5; SEQ ID 2058; 180pp; French.
XX
XX CC The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format

```

CC primers derived from the fragment of the *S. pneumoniae* genome to prime  
 CC the amplification and isolating the amplified sequences. The computer  
 CC readable medium can be used in a computer-based system for identifying  
 CC fragments of the *S. pneumoniae* genome of commercial importance, or  
 CC expression modulating fragments of the *S. pneumoniae* genome. Products  
 CC from the present invention can be used in diagnosis kits and assays, and  
 CC pharmaceutical compositions and vaccines for *S. pneumoniae*.

XX Sequence 750 BP; 195 A; 238 C; 173 G; 144 T; 0 other;

# Alignment Scores:

Pred. No.:	0.00103	Length:	750
Score:	165.50	Matches:	62
Percent Similarity:	38.95%	Conservative:	49
Best Local Similarity:	21.75%	Mismatches:	131
Query Match:	6.99%	Indels:	43
DB:	19	Gaps:	3

US-09-889-314-2 (1-496) x AAV52503 (1-750)

```

Qy 57 SerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThr 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 TCAGCATCAACAGAGTGGCTTCAGCAAGTACTGATGATGATGATGATGATGATGATGAT 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 77 GluThrAlaProGlnGlnGlyAlaAlaAlaGlyLysGluSerGluSerGluSerGluAla 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 AGTCATCAACCCATCAAGTATCTCAAGCGTGAATTCGATCAACAGAGTGCATCAAC 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 97 GlyAlaAspThrGlyAlaSerGlyAlaAlaAlaAlaThrAlaSerAsnThrAlaThrLys 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 TCAGCATCAACAGAGTGGCTTCAGCGTCACGATCAACAGTGCATGCTCAAGCAAC 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 117 IleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGluSerThrLysGluSer 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AGTCGTCGCGTTCATCAACGATGCTGATCAACCTCAAGCAAGTATCTCAAGCGTCAA 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 137 LeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaAlaAlaLeu 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 TCGGCATCAACAGAGTGGCTTCAGCATCAACAGT-----GCCTCAGCAACTACTAGTCGCGCT 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 157 SerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyVal 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 TCAGCATCAACAGAGTGGCTTCAGCATCAACAGT-----GCCTCAGCAACTACTAGTCGCGCT 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 177 ThrProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 -----GCCTCAGCGCTCAAGCAAGTATCTCAAGCG 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 197 GlyGluAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGln 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 TCTGAATTCGATCAACAGAGTGGCTTCAGCGTCACGATCAAGTACTAGTCATCAAGCA 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 217 ThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGluTyr 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 TCACAG-----GCCTCAGCGCTCAAGCAAGTATCTCAAGCG 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 237 GlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 -----AGTCATCGCGCTTCAGCAAGTACCAAGCGCTTACGTCAGCAAGCCAGT 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 GCGTCAGCGTCAAGCAAGTACCAAGCGCTTCAGCGTCACGATCAAGTACTAGTCATCAAG 549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGlyAlaAla 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 GCAAGTACCAAGTACCGCTTCAGCGTCAAGTACGATGCTGCTTCAGCAAGTACCTCA 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 GCGTCGATCAAGTACCAAGTACCGTTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 317 AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAla 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

Db 670 GCAAGTATCTCAAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 337 AlaIleLysAlaAla 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 GCCTCAGCATCAACG 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

## RESULT 43

AB067197 standard; DNA; 1163020 BP.

AB067197;

29-AUG-2002 (first entry)

Listeria innocua contlg DNA sequence #10.

Antibacterial; Listeria; food contamination; mutational analysis; infection; ds.

Listeria innocua.

WO200228891-A2.

11-APR-2002.

04-OCT-2001; 2001MC-FR03061.

04-OCT-2000; 2000FR-0012697.

(CNRS ) INST PASTEUR.

(CNRS ) CNRS CENT NAT RECH SCI.

Kunst F, Glaser P;

WPI; 2002-332479/37.

New genomic sequences from *Listeria* species, useful for detection, treatment and prevention of infection, also related polypeptides, antibodies and modulators

Claim 5; SEQ ID 10; 180pp; French.

The present invention relates to nucleic acid sequences

(AB067188-AB071212) from *Listeria* sp. The sequences are useful as probes

and primers for identification and/or detection of *Listeria* (e.g. as

contaminants in foods, or mutational analysis) and for analysis of

gene expression. Proteins encoded by the nucleic acid sequences can be

used to screen for compounds that modulate gene expression, replication

and pathogenicity of *Listeria* (potential therapeutic agents), also for

treating infections by *Listeria*, and are useful as immunogens in

anti-*Listeria* vaccines.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic format

directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.

Sequence 1163020 BP; 388339 A; 197589 C; 235612 G; 341474 T; 6 other;

Alignment Scores:

Pred. No.:	6.17	Length:	1163020
Score:	165.50	Matches:	125
Percent Similarity:	39.69%	Conservative:	77
Best Local Similarity:	24.56%	Mismatches:	206
Query Match:	6.99%	Indels:	101
DB:	24	Gaps:	21

US-09-889-314-2 (1-496) x AB067197 (1-1163020)

```

Qy 32 GlnGlnAspLysLeuSerGlyAsnGluThrLysGln----- 43
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18546 AAGCGAATTAACCTTCGCGGAAAGAACAGAAAAAGAAAGCAATTAATTAAGTGA 18605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 44 -----11eGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

Db 1807 TCCACAGGCTCGATTATTAAATCCGACGCTGCCATGCT-----GGAAGACAGTCTGCA 1754
Oy 109 ThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlu---GluAlaSer 127
Db 1753 ACTGCAACTTCTGTTGTAACCTTCGACAGCGGTGCAACTGGCGTTTCAATGGCAGCTTCT 1694
Oy 128 LysSerMetGlnSerThrLeuGlnSerLeu---GlnSerLeuSerAlaAlaGlnMetLys 146
Db 1693 GATAGTACAGAAACATCATGATGATTCCTAGGACACTGCCGAGGTGGTGGCTGCTCG 1634
Oy 147 GluValIleGluAlaValAla-----ValAlaAlaLeuSerGlyLysSer 160
Db 1633 GGAACAGAGCGCCGTGGTGGATAGCAACAGAGGAGCGAGCGGACAGCGACTGCA 1574
Oy 161 SerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProArgSer 180
Db 1573 GCGATGGATCAGAG-----GCTCG 1553
Oy 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200
Db 1552 GCGATCAGACAGGACAACTTGTGTAATCTAAAGCTGCTATGCTGCTGCTGATGCT 1493
Oy 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 1492 GGTGCTGTGCTCTCAGAGTTCTTCAGCCACTACAGCTGTGTACAGTTCCGACACCA 1433
Oy 221 Gly-LeuGluLysGlnAlaIleLysIleAspLysGluArgGluGluGlnGlnMet-- 239
Db 1432 GGAGGAGAACAGTGGGCTA-----AGTCACAGGAGTTTCACAGCAGCA 1385
Oy 240 -----LysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAs 257
Db 1384 CTAGAACGCGGTGATGAGCATCAGCGAGGTATTCG-GGAACAGCAGATGAGTGC 1326
Oy 257 nThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIlePhe 277
Db 1325 ATCCGATCCGTTACTGCTGATGTCAGGAAGT-----GGTGTGCTGGC 1281
Oy 277 eThrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaIleValAlaGly-----Al 294
Db 1280 AACGACTGGCGGTGATGATTCGCGCCACAGAGGTACACAGGGGTGGTGTCTC 1221
Oy 294 AlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaIleThrValAlaIleThrGlnIle 314
Db 1220 AGCAACAGCAACTGAAGCAACTTTCGGGTTTCTGCACACAGGTGTCGCAATGTTTC 1161
Oy 314 rValGln---AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGln 333
Db 1160 TGGCTCTGGGCTGGAGTAGAGGTGACCAACAGTATTCGCAACTTGGGACTC 1101
Oy 333 alleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLe 353
Db 1100 GGCAACGACTGAGCAAGGCGTGGCTT---TCTGTGCACTGTTTACAGTGCAGCAC 1044
Oy 353 uValLysAlaIleAlaLysAlaIleSerLysGlyLysSerValPheAlaLysGlyTh 373
Db 1043 TACCGGGGCTCAGTCAAGCACTTACAGAGCAGCACTCGGCTCCACAGCTGGTGC 984
Oy 373 rGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLys 393
Db 983 A-----ACCGACAC 975
Oy 393 sThrP-----ValThrValGlyValAlaValAlaAlaProAl 407
Db 974 ATGGGGGAGAGCTGACGCCGAGCAACAGCTGGGTTTCAGTACGCCGGTGGAGCAACGAC 915
Oy 407 aleuGlyLysGlyIle-MetGlnMetGlnLeu----- 417
Db 914 CACTGGCTCAGCTGACGTAGTGGTTTCCTGCACAGCAAGCACTCGGGATGAGAG 855
Oy 418 -----SerGluMetGlnGlnAsnValAlaGlnPheGln- 428
Db 854 TGCACAGCACTCCGGTATGTGCAACCGAATCTCAGCAACAGGTGTGGAGCTGGAGT 795

```

```

Oy 429 -----LysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnP 447
Db 794 CTCGGCGACAGCAGGGGGCACTGCGAGAGCAACGGCAGCGCCAGAACTGGGGCAACAC 735
Oy 447 heTrpGlnGln-----A 451
Db 734 AGGAACACACACTGTGTACAGCAGTACAGTCTGCAACGGGAACCGTGTGTAGCGGG 675
Oy 451 laseThrLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrL 471
Db 674 AACCAAGCTGGTGTGGGATC-ACAGGCGAAACACAGCGGAGTCAAG-----GACGCG 622
Oy 471 yLysGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIleAlaGlnAla 491
Db 621 GAGTGGGGGACAGAGGACACACAGGTGTGTCGCTATACAGCGCCAGCTACAGGAATTG 562
Oy 491 lLysThr 493
Db 561 GCGCAACA 554

RESULT 42
AAV52503
ID AAV52503 standard; DNA; 750 BP.
AC AAV52503:
XX 23-OCT-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:370.
XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
OS Streptococcus pneumoniae.
XX WO9818931-A2.
XX 07-MAY-1998.
XX 30-OCT-1997; 97WO-US19588.
XX 31-OCT-1996; 96US-0029960.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX Kunsch CA, Rosen CA;
XX WPI; 1998-272225/24.
XX
XX Computer-readable medium with recorded Streptococcus pneumoniae
XX polynucleotide sequences - useful in diagnostic kits and assays, and
XX pharmaceutical compositions and vaccines for Streptococcus
XX pneumoniae
XX
XX Claim 1; Page 1384-1385; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
XX the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
XX recorded on it, or a representative fragment or a sequence at least 95%
XX identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
XX SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
XX Streptococcus pneumoniae. The present invention also describes an
XX isolated nucleic acid molecule encoding a homologue of any of the
XX fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
XX nucleic acid molecule is produced by a process comprising: (a) screening
XX a genomic DNA library using as a probe a target sequence defined by any
XX of the sequences in SEQ ID NO:1 to 391, identifying members of the
XX library which contain sequences that hybridise to the target sequence and
XX isolating the nucleic acid molecules from the members; or (b) isolating
XX mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
XX molecules whose nucleotide sequence is homologous to amplification

```

Qy	47	Thrtgnglnlylrsasrthglumelcluserasplahrthlealaglialasergly	66
Db	80	CGTCACCTTCAGCAACCAAGTCGGCTACTTCACATCAACCAAGTCCTGGCTCGGCA	139
Qy	67	LysasplysrthSerserthThrlYsthrGlnthralaproglnnglyalala	86
Db	140	TCACAAAGTGGCTCAGCATGCATCAACAGGTCGCTCAGCTCAGCAAGT-AC	198
Qy	87	GlylYsgluserSergluserGlnlYsalaglYlaaspthrlYsalSerglyalala	106
Db	199	TCACATCATCAGCATCAACAGTCATCAGCCACCAAGTATCTCAGCGCTGATCGCA	258
Qy	107	AlathrrthralaserasrthThralaThrlYsllealmetglnthrserrllieglnuala	126
Db	259	TCACAGATGTGATAGCATTCAGATCAACAGATGCATCGCTTCAGGTCAGCAACG	318
Qy	127	SerlyssermetgluserthLeuglSerleuglnSerleuglnSerleualaglmetlys	146
Db	319	TCAGCTTCAGCAACCAAGTCGGCTCGGCTTCACATCAACCAAGTCCTCAGCTCAGCA	378
Qy	147	GluYalglualalavalalvalalalaleuSerglylYsserserglylYsalYslu	166
Db	379	AGTATCTCAGCGCTCGATCGCATCAACAGTCGCTCAGCTCAGCAATCTAGTCGA	438
Qy	167	GlnthrProgluleuProlysProglYvalThProarSerglualllieglnlliegly	186
Db	439	TCGCGCTTCAGCAACCAAGTCGCTCGCTTCACATCAACCAAGT-----	483
Qy	187	LeuallaleualalYsalalleglnthLeuglylualathrlYsseralaleuSera	206
Db	484	---GCTTCAGCGCTCAGCAAGTATCTCAGCGCTTAATCGGCATCAACAGATCGCTCAGCC	540
Qy	207	TyrAlaserrthrglnalaglinalaspglnthrasnlyseuglyleuglnYslglnala	226
Db	541	TCAGCAATACTAGTCGATCGATCAAGATCAACG-----AGTCATCG	585
Qy	227	lleYslleasplysgluarngluYlurtYrglnlmetYslalaalaglgnlYsser	246
Db	586	GCTTCAGCAATACCAAGCGCTTCAGCTTCAGCAACCAACCAAGTCGCTCAGCAAGT	645
Qy	247	LysaspleuglnuglyThmetasPthValasnthrlamelllealalaserValala	266
Db	646	ACCAAGCGCTTCAGCTCAGCAACCAAGTCGCTCAGCTTCAGCAATACCAAGTCGCTCA	705
Qy	267	lleThrValilleSerlleValalalalalepheThrysglylYsalYslleualaglYleu	286
Db	706	GCTTCAGCGCTGACAGAGCGCTCGCTTCAGCAAGTACCTCAGCGCTGATCAGCATCA	755
Qy	287	AlaaleglYalalaalaglYlaalalaalaglYlYsalYsalYsalYsalYsalYsal	306
Db	766	ACGATCATCAGCTTCAGTCAGTCAACAAAGTCCTTCAGCTTCAGCAATACCAAGTCGCTCG	825
Qy	307	ThlrthralathrlthrglnlleThrValglnalalavalYalglnalavallYsglnalaval	326
Db	826	GCTTCAGCATCAACGAGTGTTCAGTTCAGCTTCAGCAACCAAGTCGCTTCAGTCAAGTCATCA	885
Qy	327	lleThrthralalArGlnalalaleThThralalalelYsalYsalYsalYsalYsalYsal	344
Db	886	ACAAAGTGGCTCGGCTTCAGCAACCAAGTCGCTTCAGCTTCAGCAATACTAGT	939
RESULT 41			
ABLI9629/c			
ID ABLI9629 standard; DNA: 2708 BP.			
AC	ABLI9629;		
XX	26-MAR-2002 (first entry)		
XX	Drosophila melanogaster genomic polynucleotide seq ID NO 10360.		
XX	Drosophila; developmental biology; cell signalling; insecticide;		
XX	pharmaceutical; gene; ds.		
XX			

OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PE	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-061415O.
XX	
PA	(PEKE ) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Claim 1; SEQ ID NO 10360; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABRL6176-ABRL30511), expressed DNA
CC	sequences (ABRL0840-ABRL6175) and the encoded proteins
CC	(ABBS7737-ABBS72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from wipr
CC	at ftp.wipo.int/pub/published_pot_sequences.
XX	
SQ	Sequence 2708 BP; 635 A; 852 C; 606 G; 615 T; 0 other:
	Alignment Scores:
Pred. No.:	0.00433 Length: 2708
Score:	166.00 Matches: 125
Percent Similarity:	34.93% Conservative: 79
Best Local Similarity:	21.40% Mismatches: 236
Query Match:	7.01% Indels: 146
DB:	Gaps: 22
US-09-889-314-2 (1-496) x ABL19629 (1-2708)	
QY	6 IleserSerSerSerglyProAsPasnglnLysAsnIleMetSerglnValLeuthrSer 25
	:        :::::
Db	2155 ATTTGCGAGTCTGGACGCCG-----TCATTATGTTTGCACAATCTGGTTC 2108
QY	26 ThrProGlnglValIProGlnglnAsplyLseuSerglVasnglUthrlys----- 42
	:    :    :    :    :
Db	2107 GTATTCTTGGAATCGATGTTGAAGTTCTTGTCACGCCGACTCTCACCATCACCCT 2048
QY	43 -----GlnIleGlnGlnThrArgnglngly 50
	:::
Db	2047 TCCTGGATTACGACAGGTGCCTCACAGTAATGGTTCTGTAAGGGGTGCTAGAGCTTGGT 1988
QY	51 LysasnThrGlumetGluSerAspAlaThrIle-----AlaGlyAlaSerGlylys--- 67
	:::
Db	1987 TCCGCTCAGTAGTAGCGCTCTCGAGCTGTAATTGGAGCAGACACCACTCTCGAAGAGCA 1928
QY	68 -----AsplysThr 70
Db	1927 ACTGCTTCATTGGCGCGCGATAGAGCTCTGTAATGGTGAAGAGCTGATCTGCA 1868
QY	71 SerSerThrThrIlysthrGlnUthr-----AlaProGlnglnglValAlaAlaGlylys 88
	:::          :    :    :
Db	1867 GCTTACACAGCTGCCACCGAGATGTAAGTTGTCTCAATAAAATCCGGAATAGAGAGCTGGTGT 1808
QY	89 GlusertGluSerGlnLysAlaGlyAlaAspThrGlyValserGlyAlaAlaAlaThr 108
	:::          :



Qy	112	AsnThrAlaThrIysIleAlaMetGlnThrSerIleGluGlnAlaSer-----LysSer	129
Db	6895	AAATTTGCAAAATTAAGSTA-----GAAACAGATATCGATCATAGACTTTCACACTCTTCT	6548
Qy	130	MetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaIleGlnMetLysGluValGlu	149
Db	6949	ACGTGATGTAATGCTAAAGAGCTGTAAATTAATAATTTGCAATTCCTTGACAGCGCGTGAGGA	7008
Qy	150	AlaValValValAlaAlaLeuSerGly-----LysSerSerGly	162
Db	7009	GCCCGACGCTTCCACACAGATTACCGACAGCGTTCCTGTTAACACTATAAATAGTCTGTG	7068
Qy	163	SerAlaLysLeuGln-----ThrProGluLeuProLysProGlyValThr	177
Db	7069	ATAGCTCGAGTTACCAATACATACCTGATTGACTTCCGTCACGAAAGAAAGTAATGTACG	7128
Qy	178	ProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGly	197
Db	7129	GCAAAACAGCAAAA-----AATATTAGCAAAACAGACGAAATATCGAAGATGGA	7179
Qy	198	GluAlaThrIysSerAla-----LeuSerAsnTyrlaSerThrGlnAlaGlnAla	214
Db	7180	GGACACGCAATCGGACCAATGCTCTTGTAATTAATTTTGGACAC-----GCTGTAGA	7233
Qy	215	AspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlu	234
Db	7234	GATGAAAAAAATTCGTGAGAAAGAAAGACAGAAAGTTTAAAACTTTAGACGAAGTTAAC	7293
Qy	235	GluTyrlGlnGluMetLysAlaIleGlnGlnLysSerLysAspLeu-----	249
Db	7294	AAAGACACAAGTAAAAAAGTAAATGATGCTACGAAAAAAATCTTCAATCAGCAGATTT	7353
Qy	250	-----GluGlyThrMet	253
Db	7354	TCTACAGAAAGTACTTCTGTAAAAACGGATAGAGAGATCTACGGAGAAAGATTAAA	7413
Qy	254	AspThrValAsnThrVal---MetIleAlaValSerValAlaIleThrValIleSerIle	272
Db	7414	GCCATTGTGAAGACTTCTGATATTATTGTGAAAAAATGATGATTATACACAGAGACAG	7473
Qy	273	ValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaVal	292
Db	7474	AATATATCACTTCTACTGCTGCGTTTGGAACGACGACGCTTCCTCC-----	7521
Qy	293	GlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGln	312
Db	7522	-----GCATCGAGACAGCTGGCGATTACAAATTAATAAAGAAAT	7560
Qy	313	IleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGln	332
Db	7561	TCCGGAGTTACTGTTGAAAATCTTTTGTGAAGACGCTGAAAAGTAATGTTAGATCG	7620
Qy	333	AlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThr	352
Db	7621	GATATTACAGAAATGTTGCTTTAACGACATACAGAGTCTCTGTGAGACATTG---GGA	7677
Qy	353	LeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGly	372
Db	7678	AATGAGAGCTGCCCTATGCGAATTAAATTCATATGGAAGATCAAAATATCTGTTAAAT	7737
Qy	373	ThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSer	392
Db	7738	TTCTAATGCTTAATGAAAAAATATTTGATCTATTGTAAAGATTAATCGGAATTTGAGACG	7797
Qy	393	LysTrpValThrValGlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIle	412
Db	7798	GAAGCAAAAGGATTAAACCTGAGGAGCGGTAGCTCCGGAACCATTAATCTCAAAACCAAG	7857
Qy	413	MetGlnMetGlnLeuSerGlnMetGlnAlaAsnValAlaGlnPheGlnLysGluValGly	432
Db	7858	AATGAAATTCAT---TCAAGAGTTGAATTTGAGAAGGATTTTTCATATGAAGA-----	7908
Qy	433	LysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaSer	452

Db 7909 -----AAT 7911  
Oy 453 Lys11Leu1Aser1ySg1nhr1g1yGlu1Ser1Asn1Glu1Met1r1n1yS1Ala1Thr1y1Leu 472  
Db 7912 AGATTACTAGTCCCTCTCTAAAGCAATTGGCAAGAAATTCATGTCTCAAGTGGAAAGAA 7971  
Oy 473 Gly1Lacl1n1Leu1yS1Ala1y1r1Ala1a1le1Ser1y1Ala1le1a1a1G1y1Ala 490  
Db 7972 AACAGAGTGACTGCTGAATCTCAGAGAGCTTCTGTAGGACAGTAGACAGGAGCA 8025  
RESULT 39  
ID AAV52396 standard; DNA; 3744 BP.  
AAV52396:  
AAV52396:  
23-OCT-1998 (first entry)  
Streptococcus pneumoniae genome fragment SEQ ID NO:263.  
Streptococcus pneumoniae: S. pneumoniae; genome; diagnosis; assay;  
computer readable medium; vaccine; pharmaceutical composition; ds.  
Streptococcus pneumoniae.  
MO9818931-AZ.  
07-MAY-1998.  
30-OCT-1997; 97MO-US19588.  
31-OCT-1996; 96US-0029960.  
(HUMA-) HUMAN GENOME SCI INC.  
Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;  
Kunsch CA, Rosen CA;  
WPI: 1998-272225/24.  
Computer-readable medium with recorded Streptococcus pneumoniae  
polynucleotide sequences - useful in diagnostic kits and assays, and  
pharmaceutical compositions and vaccines for Streptococcus  
pneumoniae  
Claim 1; Page 1289-1291; 1409pp; English.  
The present invention describes a computer readable medium which has  
the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)  
recorded on it, or a representative fragment or a sequence at least 95%  
identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in  
SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from  
Streptococcus pneumoniae. The present invention also describes an  
isolated nucleic acid molecule encoding a homologue of any of the  
fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the  
nucleic acid molecule is produced by a process comprising: (a) screening  
a genomic DNA library using as a probe a target sequence defined by any  
of the sequences in SEQ ID NO:1 to 391, identifying members of the  
library which contain sequences that hybridise to the target sequence and  
isolating the nucleic acid molecules from the members; or (b) isolating  
cRNA, DNA or cDNA produced from an organism, amplifying nucleic acid  
molecules whose nucleotide sequence is homologous to amplification  
primers derived from the fragment of the S. pneumoniae genome to prime  
the amplification and isolating the amplified sequences. The computer  
readable medium can be used in a computer-based system for identifying  
fragments of the S. pneumoniae genome of commercial importance, or  
expression modulating fragments of the S. pneumoniae genome. Products  
from the present invention can be used in diagnostic kits and assays, and  
pharmaceutical compositions and vaccines for S. pneumoniae.  
Sequence 3744 BP; 1119 A; 804 C; 773 G; 1048 T; 0 other;

Db	3880	TCFACAAGCTGCAACATCATTTAAGTGAATCCACAAGGTGAGCAAGTCCAAAT-----	3932
Qy	209	SerThrglnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGln-----	225
Db	3934	TCACTAAGTGCATCAATGTGCAGAACGGTATTCACGTACAGAAATAACAGATCATTA	3993
Qy	226	-----AlaIleLysIleAspLysGluArgGluGlnIleThrGlnIleMetLysAla	241
Db	3994	AGTGCATGCACATCATTAAGCGGTTCAACAGTAGAGTGAAATCCGGCTCAACATCATCA	4053
Qy	242	AlaGlnGlnLys-----SerLysAspLeuGlnGlyThrMetAspThrValAsn	257
Db	4054	AGTAAAGTAATATCCGATTCAATCATCAATGAGCGCTGAGATGTCCCATCAACATCAGAGA	4113
Qy	258	ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePhe	277
Db	4114	AGTCAAGTGTCTCAACATCAGAAAGTTGTGTGACTCAACAGATACATCATTCGTCACTG	4173
Qy	278	ThrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAla	297
Db	4174	AGTGCCTCAATGAATCAAAATGGGTGTGAGCTCAACATCGCAAGCCAAAGTCCCTCAAC	4233
Qy	298	GlyAlaValAlaGlyAlaAlaAlaAlaAlaThrThrValAlaThr-----GlnIleThrValGln	316
Db	4234	TCACAACAGTACACAGCAGACGAAATCCGATTCCACAAGCACATCATCATATCAAGTAG	4293
Qy	317	AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla	336
Db	4294	TCACAAGCCAAAGTGAATCCACATCGACATCAACCTCACTAAAGCATTTCA---ACAAGT	4350
Qy	337	AlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAla	356
Db	4351	ATATCTAAAGTATCAACAGCCATCAGGTGTCAACAGCACATCTGCATCATTAAGTGTGTCA	4410
Qy	357	IleAlaLysAlaIleSerLysGlyIleSerLysValIlePheAlaLysGlyThrGlnMetIle	376
Db	4411	GAAGCGAATCCGATTCCACAAGCATATCCACAAGCACAGCAAGTAGCA-----	4458
Qy	377	AlaLysAsnPheProLysIleSerLysValIleSerSerLeuThrSerLysTrpValThr	396
Db	4459	-----AAATCAGAAAGTACATCCACATCCATCGCTCAGTACATCAACAGTACA	4503
Qy	397	ValGlyValGlyValValValAlaAlaProAlaLeuGlyLys-----	410
Db	4504	AGCAACTGACGATCAGCAAGCACTAGCTCAACATTCCTCAGTACAGCAAGCCCAAGTAA	4563
Qy	411	-----GlyIleMetGlnMetGlnLeuSerGlnMetGlnLysVal	424
Db	4564	TCCGATTCGTCGTACACATCATTTAAGTATGATTCAAGCTCAGCGTCATGCAAGCAGTGAA	4623
Qy	425	AlaGlnPheGlnLys-----GluValGlyLysLeuGlnAlaAlaAlaAspMetIle	441
Db	4624	TCCGATTCACAAAGCACATCAGCTCATTAAGTAATTCACAAAGTACATCAACATCATTT	4683
Qy	442	SerMetPheThrGlnPheThrGlnAlaSerLysIleAlaSerLysGlnThrGlyGlu	461
Db	4684	CGAATGTCGACT-----ATTGCAAGTACGTAGTATCGGAAAGTACATCAGAA	4731
Qy	462	SerAsnGlnMetThrGlnLysAlaThrLys	471
Db	4732	TCCTGTTCAACATCAGAAACTACAAAGTAA	4761
RESULT 38			
ID	AA167640	standard; DNA; 9726 BP.	
AC	AA167640:		
XX	27-FEB-2002	(first entry)	
DT			
DE	F. necrophorum leukotoxin protein encoding DNA.		
XX	Leukotoxin; infection; immunisation; liver abscess; foot rot;		

[illegible]



```
D 8699 GGAGCATG--GGAATGAGAGCTGCTATGAGAAATTAATTCATGAAATCAAT 8755
OY 368 ValPheAlaIysGlyThrGlnMetIleAlaIysAsnPhleProIysLeuSerIysValIle 387
DB 8756 ATCAGTATTAATAATTCATGATATGAGAAAATAATGATGTTTGTAAAGATATA 8815
OY 388 SerSerLeuThrSerIysThrValThrValGlyValGlyValValAlaAlaIleProAla 407
DB 8816 TCGGATTTGAGACGCGAAGCAAAAGCATTAACCGTAGAGCGCTACCTCCGAGCCATT 8875
OY 408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnIleAsnValAlaGlnIle 427
DB 8876 ATCTCAAAAGCAAAAGATGAATGAAT--TCAGAGGTGAAATTAGAGAGATATTTC 8932
OY 428 GlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPhe 447
DB 8933 AATGAAACA----- 8941
OY 448 TrpGlnGlnAlaSerIysIleAlaSerIysGlnThrGlyGlnSerAsnGlnMetThrGln 467
DB 8942 -----AATGAGTACTAGCCCTTCTTAAGGAATTCGAAAGAAATCAATGTC 8989
OY 468 LysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIle 487
DB 8990 AAGGTGAAAAGAAAACAGAGTGCATGCTGAATCTCAAGAGCTTCTGTAGAGCAGTA 9049
OY 488 AlaGlyAla 490
DB 9050 GCAAGGCGCA 9058

RESULT 37
AAS54978
ID AAS54978 standard; DNA: 7035 BP.
XX
XX AAS54978:
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #1290.
XX
XX Antisense: ds; prokaryotic cellular proliferation gene;
XX
XX antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX
XX 23-MAY-2000; 2000US-206848P.
XX
XX 26-MAY-2000; 2000US-207727P.
XX
XX 23-OCT-2000; 2000US-242578P.
XX
XX 27-NOV-2000; 2000US-253625P.
XX
XX 22-DEC-2000; 2000US-257931P.
XX
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX
XX P-PSDB; AAU37119.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 8615; 51pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
```

```
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
```

Sequence 7035 BP; 2534 A; 1382 C; 1316 G; 1803 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.0104	Length:	7035
Score:	167.50	Matches:	101
Percent Similarity:	39.02%	Conservative:	98
Best Local Similarity:	19.80%	Mismatches:	234
Query Match:	7.07%	Indels:	77
DB:	23	Gaps:	17

US-09-889-314-2 (1-496) x AAS54978 (1-7035)

```
OY 5 SerIleSerSerSerSerGlyProAspAsnGlnIysAsnIleMetSerGlnValLeuThr 24
DB 3334 AGTGATGATGCAATCAGCTTCACTGATCAGTGAATCAATATCAAAAGTATGACACA 3393
OY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrIleGlnIle 44
DB 3394 AGTACATCTGCTTGGTAAGTACTGACATCTGTAAGTAAAGCAATTAATCAGAACGC-- 3450
OY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
DB 3451 -----ACATCAACATCTATG--AGTCATTCACCAAGCTTAAGTACA 3489
OY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnIle 84
DB 3490 TCAGAGTCTGATTCACAGCAGCACTCAACGCGATTCATCAAGTGAAGCGATA 3549
OY 85 AlaAlaGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlyValSerGly 104
DB 3550 ---TCTGTTTCAGAGAT-----ACGCTATATCATTTA 3579
OY 105 AlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlu 124
DB 3580 AGTAATCAATATGATACAGATGATTCAGAGTCAAGTCAAGTCTGTTTAAAGTAA 3639
OY 125 GluAlaSerLysSer---MetGlnSerThrLeuGlnSerLeu-----Gln 138
DB 3640 TCATTGATGCAAAAGCATCCGATCAACATCAAGTCAAGTCAAGTCAAGCAGAT 3699
OY 139 SerLeuSerAlaAlaGlnMetLysGluValGlnValValValAlaAlaLeuSerGly 158
DB 3700 AGTACGCTTTATATGATGATGATTCGATCTGATGACACTTCACATCATTAAGTAAT 3759
OY 159 LysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProIysProGlyVal---Thr 177
DB 3760 TCACACAGTGTAGTGGCTCATTTCAACATCGAACAGTGGTAGCGCATCAACATCACG 3819
OY 178 ProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln----- 194
DB 3820 GTTAAGACGAGAGTGTCTCAACATCATTTGAGCAGATCAACAGATCAAGTTGTCTGAC 3879
OY 195 -----ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyrAla 208
```



Db 4863 CAATGAATTCGATTCGATGTAGTGAATCCAGAGAAACGATTCGAATTCGCG 4922  
 QY 434 ucllnalalalaaaspmetlleSerMetPhePhrGlnPheTpgGlnGlnAlseLys11 454  
 Db 4923 TGAGCGACGTGGAATGCAACGTCTAGTATGACCAAT----- 4959  
 QY 454 eaLaserLysGlnThrGlyLysSerAsnGluMetPhrGlnLysAlaThrLysLeuGlyAl 474  
 Db 4960 ----AGCAGACACCACTGGCAGATGAACATCAACATATCTACGATTCGAAGCCCAATGC 5015  
 QY 474 aGln 475  
 Db 5016 CAAG 5019  
 RESULT 36  
 AA167647  
 ID AA167647 standard; DNA; 11130 BP.  
 AC AA167647;  
 XX 27-FEB-2002 (first entry)  
 DT  
 XX F. necrophorum leukotoxin gene sequence.  
 DE  
 XX Leukotoxin; infection; immunisation; liver abscess; foot rot;  
 KW bactericide; vaccine; ds.  
 XX  
 OS Fusobacterium necrophorum.  
 XX WO200180886-A2.  
 PN  
 XX 01-NOV-2001.  
 PD  
 XX 25-APR-2001; 2001WO-US13240.  
 PE  
 XX 25-APR-2000; 2000US-0558257.  
 PR 24-APR-2001; 2001US-0841786.  
 XX  
 PA (UNIV ) UNIV KANSAS STATE RES FOUND.  
 XX  
 PI Nagaraja TG, Stewart GC, Narayanan SK, Chengappa KM;  
 XX  
 DR WPI; 2002-049245/06.  
 XX  
 PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing  
 PT an animal against an infection e.g. foot rot, or liver abscesses caused  
 PT by the bacterium -  
 PS  
 XX Disclosure: Page 102-107; 108pp; English.  
 PS  
 XX The invention provides an isolated Fusobacterium necrophorum leukotoxin  
 CC polypeptide and its truncated versions. The leukotoxin protein can be  
 CC expressed by standard recombinant methodology. The leukotoxin gene is  
 CC useful for preparing a vaccine which confers effective immunity against  
 CC infection caused by F. necrophorum. The vaccine is useful for immunising  
 CC an animal against liver abscesses caused by F. necrophorum and for  
 CC preventing foot rot caused by F. necrophorum infection. The present  
 CC sequence represents the F. necrophorum full-length leukotoxin gene  
 CC sequence.  
 CC  
 SQ Sequence 11130 BP; 4205 A; 1472 C; 2511 G; 2942 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0 015 Length: 11130  
 Score: 168.50 Matches: 111  
 Percent Similarity: 36.38% Conservative: 72  
 Best Local Similarity: 22.07% Mismatches: 215  
 Query Match: 7.11% Indels: 105  
 DB: 24 Gaps: 16  
 US-09-889-314-2 (1-496) x AA167647 (1-11130)  
 QY 52 AsnThrGluMetGluSerAspAlaThrIleAlaGlyAla----- 64

Db 7673 AACAGCAAAACGTTATAGTCTCTGTCGAATTCGAAATGCCGAGTCGAGTCGCAAAA 7732  
 QY 65 ---SerGlyLysAspLysThrSerSerThrLysThrGluThrAlaProGlnGlnGly 83  
 Db 7733 GGAGCTGGAATGGAGCAGCAGCAGTGCAGTTCACAAAGATGATCAACACGAGACAAGA 7792  
 QY 84 ValAlaIleAlaGlyLysGluSerSerGluSerGln----- 94  
 Db 7793 GTGAAAATTCCTAAATTAATATGACTGCAACAGTTAGATGTAATAGCAGAAATAGGATA 7852  
 QY 95 LysAlaGlyAlaAspThrGlyAlaSerGlyAlaAlaAlaThrThra----- 110  
 Db 7853 AATACAGTACTGGAATGCGTTCACCGCAGCTGGAATTCCTGCAGCCGAGTATCTGGA 7912  
 QY 111 -----SerAsnThrAlaThrLysIleAlaMetGlnThrSerLleGluGluAla 126  
 Db 7913 GTGCTTCTGTCATTAATTAATTCGAATTAAGTA-----GAAACAGATATTCATCATAGT 7966  
 QY 127 Ser-----LysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaIleGln 144  
 Db 7967 ACTTTACACTCTTCTACTGATGATTAATGTAAGCTCTTAATTAATTAATTCGATTCCTTG 8026  
 QY 145 MetLysGluValGluAlaValAlaValAlaAlaLeuSerGly----- 158  
 Db 8027 ACAGCGCGTGGAGAGAGCCGACAGTCTTCGACAGAGTACCGAGTGTCTGTTAACACT 8086  
 QY 159 ---LysSerSerGlySerAlaLysLeuGlu-----ThrProGluLeuPro 172  
 Db 8087 ATTAATAGTCTGTGATAGCTGATGTCACATAACTCTGATTTGACTTCCTGACAGAA 8146  
 QY 173 LysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAla 192  
 Db 8147 AAAGTAATGTACGCGCAAAAGAGAAAA-----AATATTAGCAAAACACGACGA 8197  
 QY 193 IleGlnThrLeuGlyGluAlaThrLysSerAla-----LeuSerAsnThyAlaSer 209  
 Db 8198 AATGCAGGAATCGAGAGCAGACCAATCGAGCCAAATGCTTGTAATTAATTTGGACAA 8257  
 QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIle 229  
 Db 8258 -----GCTGTAGAGATGTGAAAAAATTCGAAAGAAAAAGAACAGATTAAAAAACT 8311  
 QY 230 AspLysGluArgGluGluGlyGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeu 249  
 Db 8312 TTAGACGAAGTTTAACAAAGACAGATTAATAAGTAATGATGCTACGAAAAAATCTTA 8371  
 QY 249 ----- 249  
 Db 8372 CAATCAGCAGTATTTCTACAGACATCTCTGTAAGAACGCGATTAAGAGACTACTCAG 8431  
 QY 250 ---GluGlyThrMetAspThrValAsnThrVal---MetIleAlaValSerValAlaIle 267  
 Db 8432 GGAGGAAGGAATTAAGCCATTTGTGAAGACTCTGATATTTGAAAAAATCTAATAT 8491  
 QY 268 ThrValIleSerIleValAlaAlaIlePheThrGlyAlaGlyLeuAlaGlyLeuAla 287  
 Db 8492 ACAACAGAGCAGACAGAAATTAATTCACCTCTACTGCTGTTGGAACTGCAGCTCTGCT 8551  
 QY 288 AlaGlyAlaIleValAlaGlyAlaAlaIleAlaGlyAlaAlaGlyAlaAlaIleAlaThr 307  
 Db 8552 TCC-----GCATCAGAAACGTGCGACTTACA 8578  
 QY 308 ThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIle 327  
 Db 8579 AATATTAAAGAAATTCGCGACTTACTGTAATAATCTTTGTGAAGACACTGAAAAA 8638  
 QY 328 ThrAlaValArgGlnAlaIleThrAlaIleLysAlaAlaValLysSerGlyLys 347  
 Db 8639 GTAAATGTTAGATCGGATTTACAGGAAATGTTGCTTTAAACGCAATATCAAGCTCTGTA 8698  
 QY 348 AlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLys 367

OS Drosophila melanogaster.  
 XX WO200171042-A2.  
 PN  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001: 2001WO-US09231.  
 XX  
 PR 23-MAR-2000: 2000US-191637P.  
 PR 11-JUL-2000: 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 DR P-PDB: ABB59028.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 3875; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 10910 BP; 2875 A; 3179 C; 3157 G; 1699 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.0147 Length: 10910  
 Score: 168.50 Matches: 122  
 Percent Similarity: 34.69% Conservative: 66  
 Best Local Similarity: 22.51% Mismatches: 210  
 Query Match: 7.11% Indels: 145  
 DB: 23 Gaps: 20  
 US-09-889-314-2 (1-496) x ABL03131 (1-10910)  
 QY 8 SerSerSerGlyProAspAsnGlnLysAsnLleMetSerGln---ValLeuThrSerThr 26  
 DB 3607 AGTTCGGCTAATCCAGCCAAAGCGTCAGCCGCCAGCGCATTTATTCGATTCATG 3666  
 QY 27 ProGlnGlyVal-----ProGlnGlnAspLysLeuSerGlyAsnGlu 40  
 DB 3667 ACCACTGGCGTTATGACGCGATTGACGCGGACGACAAACAAAGTCGCGGTTC 3726  
 QY 41 ThrLysGlnLleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlu----- 56  
 DB 3727 TCCGCCGTGCTGAGCGAGCGTCGCGAAGCGTCCGCCGCGCTGTCGACAG 3786  
 QY 57 -----SerAspAlaThrLleAlaGlyValAserGlyLysAspLysThrSerThr 73  
 DB 3787 AAACGGCAGCTGATGCTGCCAAAGCGGCTCAAAAGTCCGCCGCAAAAGCGCGCC 3846  
 QY 74 ThrLysThrGlnThrAlaProGlnGlnGlyValAlaGlyLysGluSerSerGluSer 93  
 DB 3847 ACTAAGCTGTCGACGCA-----ACGGCTGCGGCTGAAGAGTGGCCAAAGCG 3894  
 QY 94 GlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThr 113  
 DB 3895 GCTGCACGACGCGCG-----GCGGCACGACGCACTACCGCCACAAACGCA 3939  
 QY 114 AlaThrLysLleAlaMetGlnThrSerLleGlnLleAlaSerLysSerMetGluSerThr 133

DB 3940 ACAACACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 3999  
 QY 134 LeuGlnSerLeuGlnSerLeuSerLysAlaAlaGlnMetLysGlnValGlnAlaVal 153  
 DB 4000 GCATCATCA---TCGACGGCAGTACGACGACGATTCGACGACGATTCGACGATTA 4056  
 QY 154 AlaAlaLeuSerGlyLysSerSerGlyLysAlaLysLeuGlnThrProGluLeuProLys 173  
 DB 4057 GATGCC-----GGCAAAAGGCTAGCGCCAGCAT----- 4086  
 QY 174 ProGlyValThrProArgSerGluValLleGluLleGlyLeuAlaLeuAlaLysAla 193  
 DB 4087 -----AGATGCA 4095  
 QY 194 GlnThrLeuGlnGlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAla 213  
 DB 4096 GCAACAGCTGTGACCAACGCGCAACGCAACGCAACGCAACGCAACGCAACGCAAC 4155  
 QY 214 AlaAspGlnThrAsnLysLeuGlnGlyLeuGlnAlaLleLysLleAspLysGluArg 233  
 DB 4156 GCAACAGCACCCACCGAGATCTACGCGCGCGAGGCG----- 4194  
 QY 234 GluGlnThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGly----- 251  
 DB 4195 -----AAGACCAAAATGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4242  
 QY 252 ThrMetAspThrValAlaSerThrValMetLleAlaValAlaSerValAlaLleThrVal 271  
 DB 4243 ACGGTGCGCATCGCGGTACACCGGCAACCGGAGCAAGTCCGCGC----- 4287  
 QY 272 IleValAlaLleAlaLlePheThrLysGlyAlaGlyLeuAlaGlyLeuAlaGlyAla 291  
 DB 4288 -----AGTGTGGGAGGCGCAACGCGCAACGCGCGCGCGCGCGCGCGCGCG 4323  
 QY 292 ValGlyAlaAlaAlaGly-----AGTGTGGGAGGCGCGCGCGCGCGCGCGCGCG 4383  
 DB 4324 GCGACAGCTGCGCGCAAGGCGGTGGCAACCGGAACGACGACGACGCGCGCGCGCG 4383  
 QY 304 AlaAla-----AlaThrThrValAla 310  
 DB 4384 GCTGCTAAGGAGGAGTCTGCGCGCAGATGCCAAGCATCCGTCGCGCAAACTGCTCG 4443  
 QY 311 ThrGlnLleThrValGln-AlaValValGlnAlaValLysGlnAlaValLleThrAla 330  
 DB 4444 AAGCTATCATATCGGAGGCTATATGCTATGCTGCGCAACGACGACGACGACGACG 4503  
 QY 330 IArgGln---AlaLleThrAlaAlaLleLysAlaAlaValLysSerGlyLleLysAla 349  
 DB 4504 AATGCAACGGGCGCG-TCTGCTCCAGTGCAGGGCGTTACTTGATGCTTTAAGCTTG 4562  
 QY 349 eLleLysThrLeuValLysAlaLleAlaLysAlaLleSerLysGly----- 364  
 DB 4563 CTATCAACCGTGTGATGCGCAATGTAAGCATCCACGGGTGCGATGATCTGCTGC 4622  
 QY 365 -----LleSerLysValPheAlaLysGlu 372  
 DB 4623 GAACGCTGTGCGACACACCCCGATGTTAGCAGCAACGATCATCATCATCATCATCAT 4682  
 QY 372 YThrGlnMetLleAlaLysAsnPro-----LysLeuSerLysValLleSerSerL 390  
 DB 4683 TGGCATTAATTCGTCAAGACACACGATCATCAGGCGCGCAAAATTCATCTTCAC 4742  
 QY 390 uThrSerLysThrValThrValGlyValGlyVal-----ValVal 403  
 DB 4743 GCGCGCAATGAGTGTAGTGTGCGCGGAGTAATTAATCCTAGGCCACACCGCACAC 4802  
 QY 403 AlaAlaProAlaLeuGlnLysGlyLleMetGlnMetGlnLeu----- 417  
 DB 4803 GCGCGCATTAATTCCTTGACGAGAGCTGAAGCGTCCCAAGTTGAGGTTGGAGCTGG 4862  
 QY 418 -SerGlnMetGlnLinsValAlaGlnPheGlnLys-----GluValGlyLysLe 434  
 DB 434 ----- 434

PI Ison CH, Duggan BM, Sapperstein SK;  
 XX WPI: 2002-519296/55.  
 DR P-PSDB: AAO21705.  
 XX  
 PT Human secreted proteins and polynucleotides for diagnosing, treating or  
 PT preventing disorders of cell proliferative, cardiovascular,  
 PT developmental, neurological and autoimmune/inflammatory disorders -  
 XX  
 PS Claim 5: Page 218-220; 229pp; English.

XX The invention relates to an isolated human secreted protein (SECP)  
 CC polypeptide from 63 fully defined protein sequences given in the  
 CC specification. The polypeptide is useful for the diagnosing/treating of a  
 CC disease with decreased/overexpression of SECP. Examples of disorders  
 CC associated with abnormal expression of SECP include a cell proliferative  
 CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory  
 CC disorder e.g. allergies, anaemia, asthma; cardiovascular disease e.g.  
 CC congestive heart failure, ischaemic heart disease; developmental disorder  
 CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.  
 CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.  
 CC The SECP polynucleotide and polypeptide are further useful for analysing  
 CC the proteome of a tissue or a cell type. The polynucleotide is useful for  
 CC creating knockin humanised animals (pigs) or transgenic animals (mice or  
 CC rats) to model human disease, and for somatic or germline gene therapy,  
 CC and further for generating hybridisation probes useful in mapping the  
 CC naturally occurring genomic sequence. This polynucleotide sequence  
 CC represents the DNA of a human secreted protein of the invention.

XX Sequence 6108 BP: 1313 A; 1838 C; 1971 G; 986 T; 0 other;

# Alignment Scores:

Pred. No.: 0.00442 Length: 6108  
 Score: 171.50 Matches: 108  
 Percent Similarity: 37.36% Conservative: 59  
 Best Local Similarity: 24.16% Mismatches: 187  
 Query Match: 7.24% Indels: 93  
 DB: 24 Gaps: 23

US-09-889-314-2 (1-496) x AAL39666 (1-6108)

OY 10 SerGlyProAsnThrGluMetSerGlnValLeuThrSerThrProGlnGly 29  
 DB 1714 ACAGCGAAGGACCCACCAAACT-----GTCCTAACCGGTAC----- 1752  
 OY 30 ValProGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnIleThrArgGln 49  
 DB 1733 -----AAGGACACCACTTACAGTGGGGTCCAGTGCCTGAACCTGGCCAAAGGGG 1803  
 OY 50 GlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSer----- 65  
 DB 1804 GGTGGCAAACTGGGCTGAAAGCAAGCAAAATATCGCAGACAGTACAAAGAACCTTT 1863  
 OY 66 GlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAla 85  
 DB 1864 GGCAGTGGGGTGGCTGCTGTAATGTGCCAAAGGGGCTGTCACAGACAGTGTAGAC 1923  
 OY 86 AlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyAlaSerGlyAla 105  
 DB 1924 ACAGCAAGACCGCTGACCGGACACCAAGACACATCTACTGGGCTCATGGGGCA 1983  
 OY 106 AlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrIleGlnGlu 125  
 DB 1984 GTG-----AATGTCCGCAAAAGGACGTGTCCAGACAGTGTGGACACC 2025  
 OY 126 AlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAlaGlnMet 145  
 DB 2026 ACCAAGACTGTCTTAAGTGTACCAAGACACCGCTCTCAGT----- 2067  
 OY 146 LysGluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerLysAlaLys 165  
 DB 2068 -----GGGGTACCGGCTGCTGCAATGTGGCCAAAGGGGCGTCCACAGCGGT 2115

OY 166 LeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIleGluIle 185  
 DB 2116 GTAGACACT-----ACAAAGTCTGCTCGACTGTGCACATAAGATGCTGTCCACT 2166  
 OY 186 GlyLeuAla-----LeuAlaLys---AlaIleGlnThrLeuGlyGluAlaThr 200  
 DB 2167 GGGCTCACAGGGGCTGTGACTTGTGCAAGAGGACTGTCCAGACCGCATGGACACACC 2226  
 OY 201 LysSerAlaLeuSerAsnThrLysThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
 DB 2227 AAGACTGTGTTAAGTGT-----ACCAAGATGCTGTGTGACAGTGGGCTGACCGTCTCG 2283  
 OY 221 GlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnLeuMetLys 240  
 DB 2284 AATGGGCCAAGGGGCGCTCCAG-----ACG 2310  
 OY 241 AlaAlaGluLysLysSerLysAspLeuGluGlyThrMetAspThrValAsnThr---Val 259  
 DB 2311 GGTGTAGACACGGCCCAAGACCGTGTGACCGCACCAAGGACACAGTCACTGTGGCTC 2370  
 OY 260 MetIleAlaValSerValAla-----IleThrValIleSerIleValAlaAla 275  
 DB 2371 ATGGGGCGCATGATATGTGCGCCAAAGGACCGTCCAGACCATGTGTGACACCAAGACT 2430  
 OY 276 IlePheThr-----CysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly 289  
 DB 2431 GTCCTAAGTGTGACCAAGGACACCGTGTGC---ATGGGGGTACCGGT---GCTCGAAT 2484  
 OY 290 AlaAlaValGlyAlaAlaAlaAlaGly----- 298  
 DB 2485 GTGGCCAAAGGGGCGCTCCAGGGGCGCTGGACACTACAAAGTGTCTGTGACGCACT 2544  
 OY 299 -----GlyAlaIleGlyAlaAla-----AlaAlaThrThrValAla 310  
 DB 2545 AAGACACCGCTATCCACTGGCTGCACAGGGCTGTGAACCTGGCCAAAGGACTGTCCAG 2604  
 OY 311 ThrGlnIle---ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAla 329  
 DB 2605 ACCGGCGTGGACACGACGACGACTGTCTGACCGGTACCAAGACACCGTGTGACGTGA 2664  
 OY 330 ValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysLysAlaPhe 349  
 DB 2665 GTCACTGTGCTGCTAATGTGGCC---AAAGGACCGTCCACAGAGTGTGGACACA--- 2718  
 OY 350 IleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysVal--- 368  
 DB 2719 GCCAAGACGCTGTGAGTGGCGCTTAAGATGTCAGTGTACTGTGAGTCAAGGGGCACTG 2778  
 OY 369 ---PheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIle 387  
 DB 2779 AATGTGGCCAAAGGAACCGTGCAGACCGGCGTGGAC-----GCCCTCAAGGCTGTG 2829  
 OY 388 SerSerLeuThrSerLysTrpValThrValGlyValGlyValAlaValAlaAlaProAla 407  
 DB 2830 CTTATGGTATCCAGACACACTGTCTTCACTGTGG-----GTTACCGGTCCATGAC 2880  
 OY 408 LeuGlyLysGlyLysMetGln 414  
 DB 2881 ATGCCCAAAAGGGCGTCCAG 2901  
 RESULT 35  
 ABL03131  
 ID ABL03131 standard; cDNA; 10910 BP.  
 AC ABL03131;  
 XX 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 3875.  
 DE Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX

DR WPI; 1998-272225/24.

XX Computer-readable medium with recorded Streptococcus pneumoniae

PT polynucleotide sequences - useful in diagnostic kits and assays, and

PT pharmaceutical compositions and vaccines for Streptococcus

XX pneumoniae

PS Claim 1; Page 1384; 1409pp; English.

XX

CC The present invention describes a computer readable medium which has

CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)

CC recorded on it, or a representative fragment or a sequence at least 95%

CC identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in

CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from

CC Streptococcus pneumoniae. The present invention also describes an

CC isolated nucleic acid molecule encoding a homologue of any of the

CC fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the

CC nucleic acid molecule is produced by a process comprising: (a) screening

CC a genomic DNA library using as a probe a target sequence defined by any

CC of the sequences in SEQ ID NO:1 to 391, identifying members of the

CC library which contain sequences that hybridize to the target sequence and

CC isolating the nucleic acid molecules from the members; or (b) isolating

CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid

CC molecules whose nucleotide sequence is homologous to amplification

CC primers derived from the fragment of the S. pneumoniae genome to prime

CC the amplification and isolating the amplified sequences. The computer

CC readable medium can be used in a computer-based system for identifying

CC fragments of the S. pneumoniae genome of commercial importance, or

CC expression modulating fragments of the S. pneumoniae genome. Products

CC from the present invention can be used in diagnosis kits and assays, and

CC pharmaceutical compositions and vaccines for S. pneumoniae.

XX

SQ Sequence 869 BP; 211 A; 277 C; 211 G; 169 T; 1 other;

Alignment Scores:

Pred. No.:	Length:	869
Score: 172.50	Matches: 74	
Percent Similarity: 41.14%	Conservative: 49	
Best Local Similarity: 24.75%	Mismatches: 138	
Query Match: 7.28%	Indels: 39	
DB: 19	Gaps: 6	

US-09-889-314-2 (1-496) x AAV52502 (1-869)

OY 59 AlAThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThr 78

DB 6 AGTACTAGTGCATCAGCTTCAGCATCAGCATGCGCTTCGCTCAACAGTCA 65

OY 79 ALAProGlnGlyValAlaAlaGlyLysGlySerSerGlyLysAlaGlyAla 98

DB 66 TCAGAGTCACAGTACAGTACGCTTCAGCTTCGCTTCGCTTCGCTTCAGCA 125

OY 99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIle 117

DB 126 AGCACCGAGTGGCTTCAGCAAGTACAGCTTCAGCTTCAGCTTCAGCTTCAGCG 185

OY 118 AlAlaMetGlnIleThrSerIle-----GluGluAlaSerSerSerMetGluSerThrLeu 134

DB 186 TCAGCTCAGCATCAGTATCTAGCGCTGTAATCGCATCAACGAGTGGCTTCAGCA 245

OY 135 GluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaVala 154

DB 246 AGTACTAGTGCCTTCAGCTTCAGCTTCAGCAAGTACAGCTTCAGCTTCAGCG 305

OY 155 AlAlaSerGlyLysSerGlySerAlaLysLeuGluThrProGluLeuProLysPro 174

DB 306 TCTGAATCGCATCAACGAGTGGCTTCAGCTTCAGCAAGTACTAC----- 350

OY 175 GlyAlaThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194

DB 351 -----GCCCTAGGCTTCAGCTTCAGCA 371

OY 195 ThrLeuGlyAlaIleThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAla 214

DB 372 AGTGCATCGGCTTCACATCAACGAGTGGCTTCGCTTCAGCAAGTACTAGCGCTTCAGCG 431

OY 215 ASPGlnThrAsnLysLeuGlyLysGluValAlaIleLysIleAspLysGlu----- 232

DB 432 TCAGCGTCAACAGTGCATCAGCTTCAGCTTCAGCAAGTACTAGCGCTTCAGCATCAACG 491

OY 233 ---ArgGluGluIleThrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluI 251

DB 492 AGTGCATCGGCTTCAGCAAGC-ACATCAGCTTCGTAATCTGCATCAACCAAGTGGCTCAGC 550

OY 251 yThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSe 271

DB 551 CTCAGCATGCACAGGCGCTTCAGCTTCAGCAAGTACTAGCGCTTCAGCGCTTCAGCGCTTC 596

OY 271 rIleValAlaAlaIlePheThrCysLysAlaGlyLeuAlaGlyLysAlaAlaIleAlaI 291

DB 597 -----GCCCTCAGCGCTTCAGCAAGTGCCTTCAGCAAGTACTAGCGCTTCAGCGCTTC 649

OY 291 aValGlyAlaAlaAlaAlaGlyLysAlaAlaGlyLysAlaAlaAlaAlaAlaAlaAlaAla 310

DB 650 AGCAAGTACCAAGTGCCTTCAGCTTCAGCAAGTACTAGCGCTTCAGCGCTTCAGCGCTTC 709

OY 311 -----ThrGlnIleThrValAlaGlnAlaValAlaGlnAlaValLysGlnAlaVala 326

DB 710 TGAATCGCATCAACCAAGTGCCTTCAGCTTCAGCAAGTACTAGCGCTTCAGCGCTTCAGC 769

OY 326 lIleThrAlaValAlaLysGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSer 344

DB 770 MCAGAGTGCATCGCTTCAGCATCAACCAAGTGCATCAAGTACTAGCGCTTCAGCGCTTC 824

RESULT 34

AAI39666

ID AAI39666 standard; DNA; 6108 BP.

AC AAI39666;

DT 05-SEP-2002 (first entry)

XX

DE Human secreted protein DNA SEQ ID No 110.

KW Antiartherosclerotic; cytoskeletal; HIV; antiatherogenic; antihaemic;

KW antisthmatic; cardiatic; vasotropic; neuroprotective; nootropic; SECP;

KW anticoagulant; antiparkinsonian; cerebroprotective; antiinflammatory;

KW immunosuppressive; human secreted protein; cell proliferative disorder;

KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;

KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;

KW ischaemic heart disease; congestive heart failure; neurological disorder;

KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;

KW Parkinson's disease; epilepsy; stroke; knock in humanised animal;

KW transgenic animal; gene therapy; gene; ds.

XX

OS Homo sapiens.

PN WO200238602-A2.

XX

PD 16-MAY-2002.

XX

PF 08-NOV-2001; 2001MC-US47420.

XX

XX 08-NOV-2000; 2000US-247505P.

PR 09-NOV-2000; 2000US-248642P.

PR 16-NOV-2000; 2000US-249824P.

PR 21-NOV-2000; 2000US-252824P.

PR 08-DEC-2000; 2000US-254305P.

PR 18-DEC-2000; 2000US-256448P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walla NK;

PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;

PI Ding L, Hafalia AOA, Tang YT, Bandman O, Warren BA, Honchell CP;

PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Dai PG, Tran B;



Db 3008 GTTCAGGACTTCTATGCCCTTCAAGAAATCC-----AGTCTACATGCTGCTTTAACT 3061  
 Oy 202 SerAlaLeuSerAsnThrGlnAlaGlnAlaAspGlnThrAsnLysLeu--- 220  
 Db 3062 TCCGCTTTTCCCTCGATGGGAAACAGAGCAATTTGGACAGACGGCTTATAATTAATTGANG 3121  
 Oy 221 -----GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnuAAGGluGln 235  
 Db 3122 AAGAAAAACAATTGACATTTGAAGAGCCCTGATTGTAAGAGAAAAAGAGAAAGAA 3181  
 Oy 236 TyrGlnGlu-----MetLysAlaAlaGlnLysSerLysAspLeuGlnGly 251  
 Db 3182 TTGGAGAGAGTTCGTTCTCGGCTTATCGGCTGCTCTAAAGAGAAAAAGACCTTCCTGACCCA 3241  
 Oy 252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThr----- 268  
 Db 3242 TCTCTGCCATCATCAAGAAAGAACGAGATCCCTTCAGAGCTATCGGAGCCAGCAAC 3301  
 Oy 269 -----ValIleSerIleValAla 274  
 Db 3302 CTTCATATCTCTGTCACACGAGCAATCAAGAACCTTTACTATCTACTGTTACCCACA 3361  
 Oy 275 AlaIlePheThrCysGlnAlaGlyLeuAlaGlyLeuAlaIleAlaValAlaGlyAla 294  
 Db 3362 GTGCTGGAATAATCCAGAGAGCGCTCTCCACAGCGGATCCGCGACAGAGAGAGCTCG 3421  
 Oy 295 AlaAlaAlaGlnGlyAlaAlaGlnAlaAlaAlaAlaAlaThrThrValAlaIleThr 314  
 Db 3422 TCGCGCAGCGCGGAGGAGCGCGCGCGCGCTTG----- 3457  
 Oy 315 ValGlnAlaValAlaGlnAlaValLysGlnAlaValIlePheThrAlaValArgGlnAlaIle 334  
 Db 3458 -----AGCGCCCTGCGAGAAATATGCAATGGAGGCTGTGTCATC---AAGACAGTCTC 3508  
 Oy 335 ThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaPheIleLys----- 351  
 Db 3509 AATGCCCTTGCGACAGACAAATGCGCCACCCCTCAGACGCTTCACAAACACACACAA 3568  
 Oy 352 ThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysValPheAlaLys 371  
 Db 3569 AAGATGATGAGTTCCTGCCCCCGCTCAACGGAAGCTTATGAACATGCTTAGCGAG 3638  
 Oy 372 GlyThrGlnMetIle-----AlaLysAsnPhePro-----Lys 382  
 Db 3639 ATGGAACAGCGCATGAGATGATCAAGTTTCAACAACTCACACACAGAGCTCAAG 3688  
 Oy 383 LeuSerLysValIleSerSerLeuThrSer-LysTryValThr--ValGlyValGlyVal 401  
 Db 3689 GTCAACAAGTGGCTTTGACATGACAGTCCACTGGGTGACAGCTGAGCATTTGATTAC 3748  
 Oy 402 ValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGln 421  
 Db 3749 GTAGAGAGAGCGGGGAGAGTGGTTCGGGAAGTGGCAATTAATCGCGCTCAACCGCCAG 3808  
 Oy 422 GlnAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaIleAspMetIle 441  
 Db 3809 CAGGAAACCTCAACAACATCTCAAGTGCAGCCCAATTCGCAAAAGACACACACAAACAG 3868  
 Oy 442 SerMetPheThrGlnPheThrGlnAlaSer----- 452  
 Db 3869 CAGCAGCAGCAGCAACAACACACACACATCGATGCCCGGCGCCACAGAAATTTGACAGGA 3928  
 Oy 453 -----LysIleAlaSerLysGlnThrGlyGlu-----SerAsnGlnMet 465  
 Db 3929 GAGGAGCCCGATACAGTATCAATAAAGACAGCAACCGTTGGAGGCAATTAACCG 3988  
 Oy 466 ThrGlnLys-----AlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAla 481  
 Db 3989 CCGCAGCT- AAGAGACATTATGCTTACGAGAAACCCCAATTACGAAGTGATGCC 4047  
 Oy 482 AlaIleSerGlyAlaIleAlaGlyAlaHis 491  
 Db 4048 AAGTGGCTGGAGCAGATGCAAAATCCAT 4077

RESULT 32  
 ABA09498 standard; cDNA, 3750 BP.  
 ID ABA09498  
 AC ABA09498;  
 XX 11-JAN-2002 (first entry)  
 DE Human S3-12 homologue-encoding cDNA, SEQ ID NO:1274.  
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antistatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 KW cytotatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;  
 KW antifungal; vulnery; antitumor; ss.  
 OS Homo sapiens.  
 XX MO200157188-A2.  
 XX PN 09-AUG-2001.  
 PD 05-FEB-2001; 2001WO-US03800.  
 PF 03-FEB-2000; 2000US-0496914.  
 XX PR 27-APR-2000; 2000US-0560875.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Drmanac RT;  
 XX WPI: 2001-457740/49.  
 DR P-PSDB; ABB12254.  
 XX Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -  
 PS Claim 1; Page 968-969; 1963pp; English.  
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoietic regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal

```

Db 2350 -----TCAGCCAGCAGCAGGCTTCCACAAGTGCAGTACTAGCCGCTCA 2394
OY 256 ValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAla 275
Db 2395 ATGAGTGGCTCGACGTCACGACGACGACGCTTCCACAAGTGCAGTACTAGCCGCTCA 2454
OY 276 IlePheThrCysGlyValAglLeuValAglLeuAlaAlaGlyValAlaValAglValAla 295
Db 2455 ATGAGTGGCTCGACGTCACGACGACGACGCTTCCACAAGTGCAGTACTAGCCGCTCA 2514
OY 296 ---AlaAlaGlyGlyAlaAlaValAglValAlaAlaAlaThrThrValAlaThrGlnIleThr 314
Db 2515 ACCAGTGGCTTCAACGAGTCTCGATGTCAGCAGCAGCTTCCACAAGTGCAGTACTAGCCGCTCA 2574
OY 315 ValGlnAlaValAglValAlaValAlaValAlaValAlaValIleThrAlaValAglAlaIle 334
Db 2575 ATGTCAGCGCAGCAGCAAGCTTCCACAAGTGCAGTACTAGCCGCTCA 2634
OY 335 Thr---AlaAlaIleLeuAlaValAlaValAlaValAlaValSerGlyIleLeuAlaPheIleThrLeu 353
Db 2635 ACAAGTGGCTCGACGTCACGACGACGACGCTTCCACAAGTGCAGTACTAGCCGCTCA 2694
OY 354 ValIysAlaIleAlaValAlaIleSerIysGlyIleSerIysValPheAlaValAlaValSerGlyThr 373
Db 2695 ACCAGTGGCTTCAACGAGTCTCGATGTCAGCAGCAGCTTCCACAAGTGCAGTACTAGCCGCTCA 2754
OY 374 GlnMetIleAlaValAlaValAlaValAlaValAlaValAlaValSerIleThrSerIleThrSerIys 393
Db 2755 AGCTGAGTACCTTCAATTCATCAAA---GAGAAGGTGTATTCCTTACCTTCT--- 2808
OY 394 TrpValThrValGlyValGlyValValValAlaAlaProAlaLeuGlyIleGlyIleMet 413
Db 2809 -----ACGGGTGACCAAGATTATTCGTGAAGTCTTACCTTACCTTACCTTAAAG 2862

RESULT 31
ABLO6539
ID ABLO6539 standard; cDNA; 8199 BP.
XX
AC ABLO6539;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SPQ ID NO 14099.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
XX
KM pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR P-PsDB; ABB62436.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX
PT interactions -
XX
PS Claim 1; SEQ ID NO 14099; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent

```

```

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLO1840-ABLO16175), expressed DNA
CC sequences (ABLO57737-ABLO2072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 8199 BP; 2477 A; 1928 C; 1979 G; 1815 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.00206 Length: 8199
Score: 178.00 Matches: 129
Percent Similarity: 37.58% Conservative: 101
Best Local Similarity: 21.08% Mismatches: 216
Query Match: 7.51% Indels: 167
DB: Gaps: 24

US-09-889-314-2 (1-496) x ABLO6539 (1-8199)
OY 15 GlnIysAsnIleMetSerGlnValLeuThr-----SerThrProGlnGly 29
Db 2342 CAGACCTCCACATGACAGCAAGTTCGACGCTCGGCGTCCGTCGATCCACTTCCTCGGGA 2401
OY 30 ValProGln-----GlnAspIysLeuSerGlyAsnGluThr 41
Db 2402 GTTCCCGGTCACCCCAATCAAAAGCCCTTCTTCCACAGTGTACCTGCTGGATGCT 2461
OY 42 LysGlnIleGlnGlnThrArgGlnGlyAsnThrGlnMetGluSerAspAlaThrIle 61
Db 2462 AAAAGATC-----AATGCTGCTGTGACGACGGCCCTTGGCGGA 2500
OY 62 AlaGlyAlaSerGlyLysAspIysThrSerSerThrIysThrGluThrAlaProGln 81
Db 2501 GCTGCAGCAGGTGGA-----AATGCCACTCCCTTCTGCTCTGCTCCCGGA 2548
OY 82 GlnGlyValAlaAlaGlyLysGluSerGluSerGlnIysAlaGlyAlaAspThrGly 101
Db 2549 ATGTCACGTGAGTGCAGCAACTGCTCTTCG-----GCCGGT-----GGA 2587
OY 102 ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLys----- 116
Db 2588 GTGAGCGATTGAGGCTTCCAGTCCGGCTGTGAGTGGGAAGAAGCGCAAGCGACAC 2647
OY 117 IleAlaMetGlnThrSerIleGluAlaSer----- 127
Db 2648 ATCCGATTGAGCTGAGAGCGAAGGAGGCTTAAGCTCCAGCTTACAGTCTACA 2707
OY 128 -----LysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAlaGln 144
Db 2708 ATGGCGCGAAGAGTGGAGATGATTCGAACATGAGAGCCCTTGATTCAGCGCGAA 2767
OY 145 -----MetLysGluValGluAlaVal 151
Db 2768 GGTCTGTCGTCGAGTGGCAATTCCTTCTGCTTGCAGTGGCATCAGCCGCTCAGCTCAT 2827
OY 152 ValValAlaAlaLeuSerGlyLysSerSerGlySer-----AlaLys 165
Db 2828 CAACATGCTTCACTCTCGGAGCAATCATCTGACAGATTGACAGACGACGCTCCCGGA 2887
OY 166 LeuGluThrPro-----GluLeuProLys 173
Db 2888 ATGAAGACCTCTCCCTCCATGGTGCACCAATGAGACTTTCATCAAGCCTCGCGAA 2947
OY 174 ProGlyValThrProArg-----SerGlu 181
Db 2948 ATGAATATGACTGAGATGCTTAAAGTCGGCTTCCAGTTCGGAGCCATGATTTGAGTGAG 3007
OY 182 ValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLys 201

```



DB 3266 GGATTCTCAGTGCAGGACGCTGTCAGACAGACTGCCAAGGNGCAGAGACTGCC 3325  
 QY 378 LysAsnPhrolylsLeuSerlyValIleSerSerLeuThrlyTrpValThrVal 397  
 DB 3326 AGCGAAGCGGCGAGTCCCGCATGCTCCGATGCTGCTTCAAGTCCCGCAGTCCG 3385  
 QY 398 GlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu 417  
 DB 3386 GCTGCAAGCTTTGACAAAGCTGCCAGCGCT---GCGGAAGCGCGACCTTCAGAGTCCGCG 3442  
 QY 418 SerGlnMetGlnGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 428  
 DB 3443 AGCGCGGCTCTGATTGACGCGACCTCAAGGACGACGAGTGGCGCAAGCTCCAGTCCG 3502  
 QY 429 LysGluValGlyLys-----LeuGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 440  
 DB 3503 AGCGAAGCGGCTGCTCAAGCATCACCGCAACATCTGTATGCTGTGCTGCAAGTTCGACT 3562  
 QY 441 IleSerMetPheThrGlnPheThrGlnGlnAlaSerlyIleAlaSerlyGlnThrGly 460  
 DB 3563 GCCAGTGAAGCGGCTTCAATGCTCATCAGCAGGCTCCAGTCCAGTGAAGCGACTGGC 3622  
 QY 461 GluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480  
 DB 3623 CATGCAAGTACTGCTACAGCAAGTCAAGCAAGT---GCCGATCCAGTCTGCGTCCAGGTAC 3679  
 QY 481 AlaAlaIleSerGly 485  
 DB 3680 CCAAGTGAAGTGGG 3694  
 RESULT 30  
 ABN68049  
 ID ABN68049 standard; DNA: 2910 BP.  
 XX ABN68049;  
 XX 01-JUL-2002 (first entry)  
 DE Streptococcus pneumoniae serotype 4911.  
 XX Streptococcus pneumoniae serotype 4911.  
 KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.  
 XX Streptococcus agalactiae.  
 OS Streptococcus agalactiae.  
 PN WO200234771-A2.  
 XX 02-MAY-2002.  
 PD 29-OCT-2001; 2001WO-GB04789.  
 PF 27-OCT-2000; 2000GB-0026333.  
 PR 24-NOV-2000; 2000GB-0028727.  
 PR 07-MAR-2001; 2001GB-0005640.  
 XX (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX Telford J, Masigiani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Telford J, Masigiani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Telford J, Masigiani V, Margarit Ros YI, Grandi G, Fraser C;  
 DR WPI: 2002-352536/38.  
 DR P-PSDB: ABP27418.  
 XX New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX Claim 7, Page 3556; 4525pp: English.  
 PS The invention relates to a protein (ABP25A13-ABP30895) from group B  
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN68049-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 CC  
 SQ Sequence 2910 BP; 946 A; 602 C; 600 G; 762 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.000508 Length: 2910  
 Score: 179.00 Matches: 83  
 Percent Similarity: 36.50% Conservative: 63  
 Best Local Similarity: 20.75% Mismatches: 194  
 Query Match: 7.36% Indels: 60  
 DB: 24 Gaps: 8  
 US-09-889-314-2 (1-496) x ABN68049 (1-2910)  
 QY 30 ValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThrArgGln 49  
 DB 1795 ATACTTAACCAAGTGTATGATTTCACTGTTCTGATATTAACCAAGTATGCT 1854  
 QY 50 GlyLys-----AsnThrGlnMetGluSerAspAlaThrIle----- 61  
 DB 1855 TCTAAATATTCTTATGCGCATCAAGTATGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1914  
 QY 62 -----AlaGlyAlaSerGlyLysAspLysThrSerSerThrLys 75  
 DB 1915 AAACCTTTATCAAGAAAGTGTCTTACAGAGTGGTGCAGCTCAAGCAAGTCAAGCTCC 1974  
 QY 76 ThrGluThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerLys 95  
 DB 1975 ATGAGTCTTCGACATTCAGCAAGTATGAGTGTTCATGAGTGTTCATGAGTGTTCATG 2034  
 QY 96 AlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThr 115  
 DB 2035 ACCAGTCTTCACAAAGTGCAGATGATGAGTGTTCATGAGTGTTCATGAGTGTTCATG 2094  
 QY 116 LysIleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGlu 135  
 DB 2095 ACCAGTGTCCACTACCAAGTGTTCACCTCAGCAAGTATGAGTGTTCATGAGTGTTCATG 2154  
 QY 136 SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluValValAlaAlaAla 155  
 DB 2155 ACTAGCGCTTCACAAAGTGTTCACAGTGCATCAGTGCATCAGTGCATCAGTGCATCAGT 2214  
 QY 156 LeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGly 175  
 DB 2215 ATGAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTC 2256  
 QY 176 ValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195  
 DB 2257 -----GCAAGTACTAGCGCTCAGTACGAT 2280  
 QY 196 LeuGlyGluAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAsp 215  
 DB 2281 GCCTGACCTCAGCAGCAGCAGTCTTCACAAAGTGTTCATGAGTGTTCATGAGTGTTCATG 2340  
 QY 216 GlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGlu 235  
 DB 2341 GCCTGACG----- 2349  
 QY 236 TyrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspThr 255



PD 14-FEB-2002.  
 XX 08-AUG-2001; 2001WO-NZ00160.  
 XX 08-AUG-2000; 2000US-0634238.  
 PR 28-NOV-2000; 2000US-0724623.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (VIAL-) VIALACTIA BIOSCIENCE NZ LTD.  
 XX Glenn M, Havukkala J, Bloksberg LN, Lubbers MW, Dekker J,  
 PI Christenson AC, Holland R, O'toole PW, Reid JR, Coolbear T;  
 DR WPI: 2002-241760/29.  
 P-PSDB: AAE20110.  
 XX New polynucleotides and polypeptides from *Lactobacillus rhamnosus*,  
 PT useful in e.g. improving the flavor, aroma, texture and health-related  
 PT benefits of milk-derived products, or in increasing properties of  
 PT microbes -  
 XX  
 PS Claim 2: Fig 63; 257pp; English.  
 XX  
 CC The present invention relates to a new isolated polynucleotide comprising  
 CC a sequence present in *Lactobacillus rhamnosus* strain HN001 and encoding a  
 CC polypeptide capable of modifying the flavour, aroma, texture, nutritional  
 CC and health benefits of milk-derived products, and/or survivability of  
 CC microbes in dairy manufacturing processes. The polynucleotides are useful  
 CC for improving the properties of microbes used in the manufacture of milk-  
 CC derived products such as cheeses, yogurt, fermented milk products, sour  
 CC milks and buttermilk; in modifying the flavour, aroma, texture and health  
 CC -related benefits of milk-derived products and in increasing the survival  
 CC of microbes during industrial fermentation processes. The bacteria may be  
 CC used to increase resistance to enteric pathogens and anti-infection  
 CC activity, including treatment of rotavirus infection and infantile  
 CC diarrhoea; aid in lactose digestion; as anti-cancer and anti-mutagenesis;  
 CC liver cancer reduction; reduction of small bowel bacterial overgrowth;  
 CC immune system modulation and treatment of autoimmune disorders and  
 CC allergies; treatment of allergic responses to foods; reduction of blood  
 CC lipids and prevention of heart disease; antihypertensive effect;  
 CC prevention and treatment of urogenital infections, *Helicobacter pylori*,  
 CC or hepatic encephalopathy; treatment of inflammatory bowel disorder and  
 CC irritable bowel syndrome; modulation of endocarditis; and for improved  
 CC protein and carbohydrate utilization and conversion. The transgenic  
 CC microbial population can be administered to a mammal as an anti-  
 CC carcinogenic agent. The present sequence is *Lactobacillus rhamnosus*  
 CC outer membrane protein rompa gene.  
 XX  
 SQ Sequence 4645 BP; 1279 A; 1062 C; 1261 G; 1043 T; 0 other:  
 XX  
 Alignment Scores:  
 Pred. No.: 0.000745 Length: 4645  
 Score: 180.00 Matches: 128  
 Percent Similarity: 35.588 Conservative: 73  
 Best Local Similarity: 22.65% Mismatches: 266  
 Query Match: 7.60% Indels: 98  
 DB: 24 Gaps: 20  
 US-09-889-314-2 (1-496) x AAD31881 (1-4645)  
 QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
 DB 2054 GACGGTAATTGGAGTTACAGACACTGCCGACTGGGCTTCTTACCTCCAAATGCTTCATT 2113  
 QY 21 GlnValLeuThrSer----- 25  
 DB 2114 GAAGTTATTTCAGTCTCGATTTCATTTGAAAAACGACGAGCTAGTCAACGATGCC 2173  
 QY 26 ---ThrProGlnGlyValProGlnGln-----AspLysLeuSerGlyAsnGlnLys 42  
 DB 2174 GAACGCCCAAGACGAGCATCCAGCTGCAGCTGACGACTTATCAACGCCCAATAGTGTGCT 2233  
 QY 43 GlnIleGlnGlnThrArgGlnGlnLysAsnThrGlnMetClnuSerAspAla----- 59

DB 2234 GATGTCAACAGCTTACACAGCAAAAGCTACAAAGTCTGCTGTAGTATGCGCGAGTTAT 2293  
 QY 60 -----ThrIleAlaGly-----AlaSerGlyLysAspLysThrSer 72  
 DB 2294 GCAGGTGAAGCCCAATCGATTCGTGGCAGTCACTGTATATATGCAATCAAGTCTCTTC 2353  
 QY 73 ThrThrLysThrGlnThrValProGlnGlnGlyValAlaAlaGlyLysGlnuSerGlu 92  
 DB 2354 GCCAGTATGCTGAGACAGCAATCGCAAAATGCTTGGCAGCTAGCAAGTCTCTCGGCT 2413  
 QY 93 SerGlnLysAlaGlyAlaAspThrGly---ValSerGlyAlaAlaAlaThrThrAlaSer 111  
 DB 2414 ACTTCACAGCGCGAGCGCCGCCACCAATCGCGCAAGTGGCAAGAGCGCGTATGACGCTCA 2473  
 QY 112 AsnThrAlaThrLysIleAla---MetGlnThrSerIleGlnGlnAlaSerLysSerMet 130  
 DB 2474 TCTGCAGCTGCTGCCGTATGACCTGTGATGATGACAACTCTGACCCCTCTTAT 2533  
 QY 131 GluSerThrLeuGlnuSerLeuGlnuSerLeuSerAlaAlaGlnMetLysGlnValGlnAla 150  
 DB 2534 GATTCCTACGCTTCTGAGCGCCAGTCCGCTTCTGCTGCTAATGATAGTTCGGGATATGCC 2593  
 QY 151 ValVal---ValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnuThrPro 169  
 DB 2594 ACTGCATCATTTTGACAGCAAGTTCCTGCGGCTGCCATGAGCGGAGCTTATCGACAGCG 2653  
 QY 170 GluLeuProLysProGlyValThrProArgSerGlnValIleGlnIleGlyLeuAlaLeu 189  
 DB 2654 CAAGTTGCTGCCAAGGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2713  
 QY 190 AlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsnThrAlaSer 209  
 DB 2714 GCTAGTGCAGCTCAAAAGC-----GACTCCAGAAATAAACAAGCAGCTCAGCTACTA 2764  
 QY 210 ThrGlnAlaGlnAla---AspGlnThrAsnLysLeu----- 220  
 DB 2765 GCAGAGACGACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2824  
 QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnuArgGlnuGlnuGlnuLys 240  
 DB 2825 GCGCGCAAGCTCCAGTCCAGCGCAAGCGGCTCAAGCATGCAATCATCTGCTGATGCT 2884  
 QY 241 AlaAlaGlnLysSerLysAspLeuGlnuGlyThr----- 252  
 DB 2885 AGTGTGCAAGTTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2944  
 QY 253 ---MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271  
 DB 2945 GCCAGTACGCTGCTGCGTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 3004  
 QY 272 IleValAlaAlaIleThrValLysGlyAlaGlyLysAlaGlyLysAlaIleAla 291  
 DB 3005 ---GCGGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 3061  
 QY 292 ValGlyAlaAlaAlaAlaGlyLysAlaAlaGlyLysAlaAlaAlaAlaAlaAlaAlaAla 311  
 DB 3062 AGCGAGGACGCAAGGCA-----AGCACTAAGCGAGTCCGCGCAACGAGCCGCGG--- 3112  
 QY 312 GlnIleThrValGlnAlaValAlaGlnAlaValLys---GlnAlaValIleThrAlaVal 330  
 DB 3113 -----GCCGTGCTTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 3145  
 QY 331 ArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaIlePhe 350  
 DB 3146 GAACAGCGCAAGCGCTCAAGTCCGATGTGTCGAAGAAGTCCGCGCGAGCAGCGCT 3205  
 QY 351 LysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysVal----- 368  
 DB 3206 AACAGTAATGCGAGTCCCGCAGCAGTCCGACCAAGCTGTGTGATAGCAAAACCGCAGCA 3265  
 QY 369 -----PheAlaLysGlyThrGlnMetIleAla 377

Pred. No.: 0.431 Length: 1437668  
 Score: 182.50 Matches: 123  
 Percent Similarity: 38.15% Conservative: 83  
 Best Local Similarity: 22.78% Mismatches: 219  
 Query Match: 7.70% Indels: 115  
 DB: 21 Gaps: 24

US-09-889-314-2 (1-496) x AAA81490 (1-1437668)

OY 17 AsnIleMetSerGlnValIleuThrSerThrProGlnGlyValProGlnGln----- 33  
 Db 533026 AATCTGGTGGAGCATGACGAGCCCTACGCCATTCATGCTGCCCAAGGACGATTA 533085  
 OY 34 ---AspIleuSerGlyAsn-GluThr-----LysGlnIleGlnGln 47  
 Db 533086 ACATCGAAGCGGCGGAAATGACACTATGCGGTGAGAGAGCTCACTACGACGACAAAC 533145  
 OY 47 TargGlnGlyIleuSerGlnValMetGluSerAspAlaThr-IleAlaGlyAlaSerGlyL 67  
 Db 533146 TTGACAGCCAAAGGCGCGCAT-----TCTCGCATGCTACGACGACAAAGCAC 533196  
 OY 67 yAspIleuSerSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 87  
 Db 533197 ACGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 533249  
 OY 87 IlyGlnGluSerSerGluSerGlnValIleuSerGlyValIleuSerGlyValIleuSerGly 107  
 Db 533250 -----GAATCTGCCATCTGCAATCAAGTTGGGATGCCAACTGACGACGACGACGAC 533301  
 OY 107 IaThrThrIleSerAsnThrIleThrIleThrIleThrIleThrIleThrIleThrIleThr 126  
 Db 533302 TTGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 533352  
 OY 126 IaSerIleuSerMetGluSerThrIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 145  
 Db 533353 CACGGGCGCATGCGCAAGATTATCTCGAAGGATCAAAAGCAGCATCCACACGACGACGAC 533412  
 OY 145 eTlyGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 165  
 Db 53413 TGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 53472  
 OY 165 IlyeGlnIleuThrProGlnIleuProGlnIleuProGlnIleuProGlnIleuProGln 181  
 Db 53473 CATTGCAATTCGCGAGTTTCACCGGTCGCGGTCGCGGTCGCGGTCGCGGTCGCGG 533532  
 OY 181 IuValIleGlnIleGlnIleuAla---IleuAlaIleGlnIleuGlnIleuGlnIleuGln 200  
 Db 533533 ACATTGCAATTCGCGAAGGACATCTGAAACCCAAATCGAAACCTC-----A 533583  
 OY 200 hrIlySerIleuSerAsnThrIleuSerThrGlnIleuGlnIleuGlnIleuGlnIleuGln 220  
 Db 533584 CCAAGCAGCGCGGAGTATGCTTATTTGAACAACCTCAAGTTGCGAAACATCAAC----- 533639  
 OY 220 euGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 240  
 Db 533640 -----TGGAAATCAGGTGACGCTGCTGATTAATGAGCAAAACAG----- 533684  
 OY 240 yAlaIleuGlnIleuSerIleuSerIleuSerIleuSerIleuSerIleuSerIleuSerIleu 260  
 Db 533685 -----GAGGGCATGACACCGCGGACCGGACGACGACGCTGCG 533715  
 OY 260 eTlleAlaIleuSerValIleuThr-----ValIleuSerIleuValIleuAlaIleuThr 278  
 Db 533716 TCGTTATCGTGTAAACCTATTGACCTACGCGCAGCTTCGCGCGGACCGCGGAA 533775  
 OY 278 hrCysGlyIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 298  
 Db 533776 CCGCGGCGCGCGGACGCGAGGA---GCGGAGGAGCGCGGACGAGGACGCGGAA 533832  
 OY 298 IyGly---AlaIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 317  
 Db 533833 CTGGAGTACGACGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 533892

OY 317 IaValIleGlnIleuValIleuGlnIleuValIleuThrIleuValIleuValIleuValIleu 337  
 Db 533893 CTATACACGACGCGGACGCGAAGCGGACGCTGCGCGCTC-----G 533934  
 OY 337 IaIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 352  
 Db 533935 CCAGCGAGCGCGGTTTCCCTCATCAACACGACGACGACGACGACGACGACGACGACGAC 533994  
 OY 352 hrIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 362  
 Db 533995 AACTGGCGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 534054  
 OY 362 eTlyGlnIleuSerIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 380  
 Db 534055 TGCAGGCGATTAAGGCGGCTGCAACACGACGACGACGACGACGACGACGACGACGAC 534114  
 OY 381 ---ProIleuSerIleuValIleuSerIleuSerIleuSerIleuSerIleuSerIleuSer 400  
 Db 534115 GTCCCGCAGCAGGCAAACTGACCGCTTAACCTGATCAACACGACGACGACGACGACGAC 534174  
 OY 400 IlyValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 420  
 Db 534175 ATACCGCGCATCAACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534234  
 OY 420 eGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 437  
 Db 534235 TA-----GTCAATGACGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 534285  
 OY 437 IaIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleu 457  
 Db 534286 GCGAAGCAGTCAAT-----GCCACAGATATGATGATGATGATGATGATGATGATGAT 534318  
 OY 457 yGlnIleuSerIleuSerIleuSerIleuSerIleuSerIleuSerIleuSerIleuSer 470  
 Db 534319 CCGTAGAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 534378  
 OY 471 -----LysL 472  
 Db 534379 GCGCGGCGATGCGGAGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 534438  
 OY 472 euGlnIleuGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 490  
 Db 534439 TGTCAACCGCAAGAGCGCAAAAGTATGATGATGATGATGATGATGATGATGATGAT 534494

RESULT 29  
 AAD31881 standard; DNA: 4645 BP.  
 XX AAD31881:  
 XX 18-JUN-2002 (first entry)  
 DE Lactobacillus rhamnosus outer membrane protein rompa gene.  
 XX  
 XX Enzyme: flavour; aroma; texture; nutritional; dairy manufacture; therapy;  
 XX fermentation process; anti-infection; rotavirus infection; heart disease;  
 XX infantile diarrhoea; lactose digestion; anti-cancer; autoimmune disorder;  
 XX anti-mutagenesis; immune system modulation; allergy; Helicobacter pylori;  
 XX anti-hyperemesis effect; urogenital infection; hepatic encephalopathy;  
 XX bowel syndrome; endocarditis; transgenic microbe; outer membrane protein;  
 XX rompa gene; ds.  
 XX  
 XX Lactobacillus rhamnosus HN001.  
 XX  
 XX key Location/Qualifiers  
 XX CDS 74..465  
 XX /tag="a  
 XX /product="Outer membrane protein"  
 XX /note="CDS does not include start codon"  
 XX /partial  
 XX  
 XX WO200212506-A1.

```

Db 233413 TGAGCAGCAGCAAAATCTACTCTATGCGCAAAAACAGCAGGAGGCGAGTAATCATCGAAA 233472
OY 145 ySLeuglThrProgluueuProlysProglYval-----ThrProAtgSerg 181
Db 233473 CCTTCGAATTCGCGAGTTTCACCGGTCCTGCGCCCTACTGTCGCCAGCCGGCGGTT 233532
OY 181 luValilegluileglYleuAla----LeuAlaLysAlaileglInThrleuglYguAlaT 200
Db 233533 ACATGTGATGATTTCCGAAGCAATCTGAAAACCAATGCAAAACCCCTC-----A 233583
OY 200 hrLysSerAlaLeuSerSntYrAlaSerThrGlnAlaGlnAlaSpGlnThrAsnLysL 220
Db 233584 CCAAACAGCCCGCATGATGCTGTTTGAAAACAATTCAGATTGCGAAAAACATCAAC---- 233639
OY 220 euGlYleuGlYusGlnAlaileYsIleAspLysGlnArgLunArgLunYrGlnGlnMetL 240
Db 233640 -----TGGAATCAGGAGGCGAGCTTGCTTACGATTAATGGAATGGAACAAAG---- 233684
OY 240 ySAlaAlaGlunGlnYserLysAspLeuGlYThrMetAspThrValaSnThrValM 260
Db 233685 -----GAGGCATGACACCCCGCAGCAGCAGCAGCTGTG 233715
OY 260 etIleAlaValaSerValaAlaIleThr-----ValIleSerIleValaAlaAlaIlePheT 278
Db 233716 TCGTTATCGTCTTAACCGATATGACCTACGCGCGCACTGTCGCCCGCGCAGCGCGGAA 233775
OY 278 hrcYsGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValaGlyAlaAlaAlaAlaG 298
Db 233776 CGCGCGGCGCGCGCAGCGAGGA---GCGGAGAGAGCGCCAGCAGCAAGCGAGCGGAA 233832
OY 298 lYgLY---AlaAlaGlyAlaAlaAlaAlaAlaThrThrValaAlaThrGlnIleThrValaGlnA 317
Db 233833 CTGGAAGTACACAGCAGCAAGCGGACACACCGAGATAGCAGCAGCAGCAGCAGTACGTCGAG 233892
OY 317 laValaValaGlnAlaValaLysGlnAlaValaIleThrAlaValaArgGlnAlaIleThrAlaA 337
Db 233893 CTATCACACACACCGCAGCAAGCGCAGCTGCGCATCTC-----G 233934
OY 337 laIleLysAlaAlaVala-----LysSerGlyIleLysAlaPheIleLysT 352
Db 233935 CCAGCCAAAGCCGAGTTCTCTCATCAACACAAAGAGACATTAACCTACCTCGAAAG 233994
OY 352 hrLeuValLys-----AlaIleAlaLysAlaLys 362
Db 233995 AACTGGGCAAAAGCAGCACCTGACAGAGCGCCACCGCCGTAACCGAGGCGTAC 234054
OY 362 etLYsGlyIleSerLysValaPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe---- 380
Db 234055 TGCAGGGCTTAAGCGGCGTGAACACCCAGCAGCGCAACCGCTGACAAACATTTTCACA 234114
OY 381 --ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValG 400
Db 234115 GTCCCGCAGCAGCAAACTGACCGCTGACCTGATCAACAGCACCGCTGCCGCAAGTGTCC 234174
OY 400 lYValaValaAlaAlaPheAlaLeuGlyLYsGlyIleMetGlnMetGlnLeuSerGlnM 420
Db 234175 ATACCGCATCAACAGGCGCAGCTGTAAGACACACTGGGCGATGCCGACGCGGTGCGCA 234234
OY 420 etGlnGlnAsnValaAlaGlnPheGlnLysGluValGly-----LysLeuGlnAlaA 437
Db 234235 TP-----CTCAGTACCGTACACGAGGAGTAGCGAGCAAAATTCAAATTTAATCTCA 234285
OY 437 laAlaAspMetIleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerL 457
Db 234286 GCGAAGACTACACTT-----GCCCAACAAGATAGCCCATG 234318
OY 457 ySglnThrGlyLusSerAsnGlnMetThrGlnLysAlaAlaThr----- 470
Db 234319 CCGTAGACAGGCTGTGATCGGGGTAGCAAAATTAAGGCAAAATGTGGGAGCGGCGCAATCG 234378
OY 471 -----LysL 472
Db 234379 GCGCGGCACTCGCGAGATGTGGAGAAACCTGTTGGACGAGCGCATGTAGCGCAAAC 234438

```

```

OY 472 euGlYAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIleAlaGlyAla 490
Db 234439 TGTCAACCCCAAGACGCCAAAAAGTCATAGCTACTCGCAGATTATCGCAGCGACG 234494

RESULT 28
AAA81490
ID AAA81490 standard; DNA; 1437668 BP.
XX
XX AAA81490:
AC
AC 04-DEC-2000 (first entry)
DT
DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
XX KM Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX KM Meningococcus B; MenB; ds.
XX
XX OS Neisseria meningitidis.
XX
XX PN WO200022430-A2.
XX
XX PD 20-APR-2000.
XX
XX PF 08-OCT-1999; 99WO-US23573.
XX
XX PR 09-OCT-1998; 98US-0103794.
XX PR 30-APR-1999; 99US-0132068.
XX
XX PA (CHIR ) CHIRON CORP.
XX
XX PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
XX PI Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
XX PI Rappuoli R, Pizza M;
XX DR WPI: 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisseria infections, for example, N.gonorrhoea.
XX
XX PS Claim 7: Page 866-1272; 1760p; English.
XX
XX CC The present invention describes methods of obtaining immunogenic
XX CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
XX CC represent specifically claimed Neisseria meningitidis genomic DNA
XX CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
XX CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
XX CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
XX CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX CC AAA81452 represent Neisseria meningitidis MenB polynucleotide OMF
XX CC sequences, which are all used in the exemplification of the present
XX CC invention. The nucleic acid sequences, protein sequences, and antibodies
XX CC against them, can be used in the manufacture of a composition. The
XX CC composition can be used as a medicament (or in the manufacture of a
XX CC medicament) for treating, preventing or diagnosing infection due to
XX CC Neisseria bacteria. For example, some of the identified proteins could
XX CC be components of vaccines against Meningococcus B; against all serotypes;
XX CC and/or against all pathogenic Neisseriae. Identification of sequences
XX CC from the bacterium will also facilitate production of biological probes,
XX CC particularly organism-specific probes. Attempts to make efficacious
XX CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
XX CC Multivalent vaccines have also been tried but none have successfully
XX CC overcome antigenic variability. The provision of further, complete
XX CC sequences may provide an opportunity to identify secreted or surface
XX CC exposed proteins that may be presumed targets for the immune system and
XX CC which are not antigenically variable or at least more conserved than
XX CC other more variable regions.
XX
XX SQ Sequence 1437668 BP; 344338 A; 353206 C; 385074 G; 355045 T; 5 other;

```

Alignment Scores:

```

OY 337 lalleysalaalaval-----LysSerGlyIleLysalaPheIleLysT 352
    || :|||:|||||
DB 20407 CCGAGCCAGCGCATGTTTCCCTTCATCAACAAAGAGACCAATACCTGTAAG 20348
OY 352 hrleuVallys-----AlaIleAlaLysAlaIles 362
    ||| |||
DB 20347 AACCTGGGCAAAACACACCGCTCAGACAGCGCCACCGCGCTAACCGCAGCGCTAC 20288
OY 362 erlySGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe----- 380
    ||:|||||:|||||
DB 20287 TGCAGGGCATACCGGGCTGTAACACCAACAGCCAGCCGCTCAGCAAACTTTTCACA 20228
OY 381 --ProlySerLysValIleSerSerLeuThrSerLysValThrValGlyValG 400
    ||| |||
DB 20227 GTCGCCGAGCAGCAAACTGACCTAACCTGATCAACACACCGCTGCCGCAAGTGTCC 20168
OY 400 lyValValValAlaAlaPheAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGln 420
    ||| |||
DB 20167 ATACCGCCATCAACGCGCGCAGCTGAAAGACAATTGGCGCATGCCGCTGGTGCCA 20108
OY 420 etGlnGlnAsnValAlaGlnPheGlnLysGlyValGly-----LysLeuGlnAla 437
    ||| |||
DB 20107 TA-----GTCAGTACGCTACACGAGACAGAGAGCAAAATCAATTAATCTCA 20057
OY 437 laalaaspmetIleSerMetPheThrGlnPheThrGlnAlaSerLysIleAlaSerL 457
    ||| |||
DB 20056 GCGAAGACTACATTT-----GCCACACAGATAGCCCATG 20024
OY 457 ysgInThrGlyLysSerAsnGluMetThrGlnLysAlaThr----- 470
DB 20023 CCGTAGCAGGCTGTGCTATCGCGCTAGCAATAAAGCAATATGCGGAGCGGCATCG 19964
OY 471 -----LysL 472
DB 19963 GCGGCGGAGTCGCGGAGATGGTGGAGAACCTGTGACGCGAGCGATGTAGCGAAAC 19904
OY 472 euglyAlaGlnIleLeuLysAlaIleAlaIleSerGlyAlaIleAlaGlyAla 490
DB 19903 TGTACCCCAAGAAAGCCAAAGTCAATGCTCTGACATGATTCGACAGGCGAGC 19848

RESULT 27
AAF21607
ID AAF21607 standard; DNA; 349980 BP.
AC AAF21607;
XX
XX 13-MAR-2001 (first entry)
XX
XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:108.
DE
XX
XX Neisseria meningitidis: Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
XX Neisseria meningitidis.
OS
XX
XX WO200066791-A1.
XX
XX 09-NOV-2000.
XX
XX 08-MAR-2000; 2000WO-US05928.
XX
XX 30-APR-1999; 99US-0132068.
XX 08-OCT-1999; 99WO-US23573.
XX 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Piza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani I,
XX Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappelli R,
XX Frazer CM, Grandi G;

```

```

XX DR WPI: 2000-647603/62.
XX XX
XX Neisseria meningitidis B full length genome sequence and open reading
XX PT frames are used to detect, treat and prevent Neisserial infections -
XX PS Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
XX Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
XX to AAF21613 represent fragments of the NMB genomic sequence, as the
XX sequence was too long to go in a record on its own it was split into 8
XX sequences which overlap each other at the beginning and end of each
XX sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
XX the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
XX the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
XX Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
XX AAF21606 represent PCR primers which are used in the exemplification of
XX the present invention. The NMB genome and fragments from it have
XX antibacterial activity, and can be used in vaccines and gene therapy.
XX Neisseria nucleic acids, proteins and/or antibodies which binds to the
XX proteins can be used in compositions for treating or preventing infection
XX due to Neisserial bacteria or as a diagnostic reagent for detecting the
XX presence of Neisserial bacteria or of antibodies raised to Neisserial
XX bacteria. Computers, computer memory, computer storage medium or computer
XX databases can be used in a search to identify open reading frames (ORFs)
XX or coding sequences within the NMB genome. The DNA sequences provide
XX further opportunities to find antigenic or immunogenic proteins which are
XX more effective in vaccines than the outer membrane proteins currently
XX used.
XX
XX Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.081 Length: 349980
XX Score: 182.50 Matches: 123
XX Percent Similarity: 38.15% Conservative: 83
XX Best Local Similarity: 22.78% Mismatches: 219
XX Query Match: 7.70% Indels: 115
XX DB: 21 Gaps: 24
XX
XX US-09-889-314-2 (1-496) x AAF21607 (1-349980)
OY 17 AsnIleMetSerGlnValLeuThrSerThrProGlnGlyValProGlnGln----- 33
    |||:|||||
DB 233026 AATCTGTGTCGACATGCGACGCGCTACCGCACCGCATTCATGCCCCCAAGCGCATTA 233085
OY 34 ---AspLysLeuSerGlyAsn-GluThr-----LysGlnIleGlnGln 47
    ||| |||
DB 233086 ACATCGAAGCGCGCGCGGAAATGTACACTCTATGCGCTAGAAAGAGCTCACTACGACAAAC 233145
OY 47 rArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr-IleAlaGlyAlaSerGlyL 67
    ||| |||
DB 233146 TTGACACGCCAAAGAGCGCGCAT-----TCTCGGCATCAGCTACAGCAAGAC 233196
OY 67 ysaAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyValAlaIle 87
    ||| |||
DB 233197 ACGACACACACCCACAGTCATGAAGAAC---GGCTGCCCTCAAGGGTAGTTGCA--- 233249
OY 87 lyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlnGlyValSerGlyValAla 107
    ||| |||
DB 233250 -----GAATCTGCCATATGCAATCAGGTTGGATACCAACTGCACACACAGT 233301
OY 107 laThrThrAlaSerAsnThrAlaThrThrLysIleAlaMetGlnThrSerIle---GluGln 126
    ||| |||
DB 233302 TTGAACACACACTGGGTGGCGCAAC-----ATAAGCGGAGCGGTAGGCGAGCAGG 233352
OY 126 laSerLysSerMetGlnSerThrLeuGlnSerLeuGln-----SerLeuSerAlaIleGln 145
    ||| |||
DB 233353 CACGGGCGCATGCGCAATTCCTCGAAGGATCAAAAGCATCAACACAGCAAGAACCG 233412
OY 145 etLysGlnValGlnAlaValAlaValAlaIleAlaLeuSerGlyLysSerSerGlySerAlaL 165
    ||| |||

```

DB 1492 AGCTTAACCTGCCACACACG---GGAGGAGTGTCTTCTGCT 1536  
 RESULT 26  
 ID AAA81471/C  
 XX AAA81471 standard; DNA; 49914 BP.  
 AC AAA81471;  
 XX  
 DT 04-DEC-2000 (first entry)  
 XX  
 DE N. meningitidis partial DNA sequence gnm\_19 SEQ ID NO:19.  
 XX  
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;  
 XX Meningococcus B; MenB; ds.  
 OS Neisseria meningitidis.  
 XX  
 PN WO200022430-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 08-OCT-1999; 99WO-US23573.  
 XX  
 PR 09-OCT-1998; 98US-0103794.  
 PR 30-APR-1999; 99US-0132068.  
 XX  
 PA (CHIRON) CHIRON CORP.  
 XX  
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,  
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,  
 PI Rappuoli R, Pizza M;  
 XX  
 DR WPI: 2000-318079/27.  
 XX  
 PT Isolated nucleotide sequences of Neisseria meningitidis which can be  
 PT used in the diagnosis and treatment of N. meningitidis infection and  
 PT other Neisserial infections, for example, N. gonorrhoea -  
 XX  
 PS Claim 7: Page 451-466; 1760pp; English.  
 XX  
 CC The present invention describes methods of obtaining immunogenic  
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414  
 CC represent specifically claimed Neisseria meningitidis genomic DNA  
 CC sequences. AAA81260 to AAA81303 and AAB25620 to AAB25663 represent  
 CC Neisseria DNA sequences and their corresponding proteins. AA81254 to  
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the  
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to  
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF  
 CC sequences, which are all used in the exemplification of the present  
 CC invention. The nucleic acid sequences, protein sequences, and antibodies  
 CC against them, can be used in the manufacture of a composition. The  
 CC composition can be used as a medicament (or in the manufacture of a  
 CC medicament) for treating, preventing or diagnosing infection due to  
 CC Neisserial bacteria. For example, some of the identified proteins could  
 CC be components of vaccines against Meningococcus B; against all serotypes;  
 CC and/or against all pathogenic Neisseriae. Identification of sequences  
 CC from the bacterium will also facilitate production of biological probes,  
 CC particularly organism-specific probes. Attempts to make efficacious  
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.  
 CC Multivalent vaccines have also been tried but none have successfully  
 CC overcome antigenic variability. The provision of further, complete  
 CC sequences may provide an opportunity to identify secreted or surface  
 CC exposed proteins that may be presumed targets for the immune system and  
 CC which are not antigenically variable or at least more conserved than  
 CC other more variable regions.  
 XX  
 SQ Sequence 49914 BP; 13048 A; 11994 C; 11423 G; 13448 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 0.00807 Length: 49914  
 Score: 182.50 Matches: 123  
 Percent Similarity: 38.15% Conservative: 83

Best Local Similarity: 22.78% Mismatches: 219  
 Query Match: 7.70% Indels: 115  
 DB: 21 Gaps: 24  
 US-09-889-314-2 (1-496) x AAA81471 (1-49914)  
 QY 17 AsnIleMetSerGlnValLeuThrSerThrProGlnGlyValProGlnGln----- 33  
 DB 21316 AATCTGTGGCGACGATTCAGCGCTACGCCACCGCATTTGGATGCCCCCAAGCAGCATTTA 21257  
 QY 34 ---AspLysLeuSerGlyAsn-GluThr-----LysGlnIleGlnGln 47  
 DB 21256 ACATCGAAGCCGCGGGAATTTGACACTTATCCGTAGAGAGCTCAACTACGACAAAC 21197  
 QY 47 rArgGlnGlyLysAsnThrGluMetGluSerSpsIaThr-IleAlaGlyAlaSerGlyL 67  
 DB 21196 TTGACAGCCCAAAAGGCCGACATT-----TCTGGCATTCAGCTACACGCAAGCAC 21146  
 QY 67 ysAspLysThrSerSerThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaG 87  
 DB 21145 ACGACACCCACCCACGACGATGAAACC---GCCCTGCCCTCAAGGATTTGCA----- 21093  
 QY 87 lYlsgLuserSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAla 107  
 DB 21092 -----GAATCTGCATCTGCAATCAGTTGGATTCAACTGCAGACGACACAGT 21041  
 QY 107 IaThrThrAlaSerSerThrAlaThrLysIleAlaMetClnThrSerIle---GluGlu 126  
 DB 21040 TTGAACACCACTGGGTGGCGAACCC-----ATACGGCAGGCGTAGCGGAGCAGG 20990  
 QY 126 IaSerLysSerMetGluSerThrLysLeuGluSerLeuGln---SerLeuSerAlaAlaGln 145  
 DB 20989 CACGGCCGATCCCAAGATTAATCTTCGAAGGATCAAAAAGCAGCATCCACACGAAACCG 20930  
 QY 145 eLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSerGlySerAlaL 165  
 DB 20929 TCACGACGACGAATCTCTATGCAAAACAGCAGGCGGACGATATCATCGAAA 20870  
 QY 165 yLeuGlnThrProGluLeuProLysProGlyVal-----ThrProArgSerG 181  
 DB 20869 CCTTGCATTTGCCGAGTTTCACCGGTCCGTTGCCGCCGCTACTGTCGCCGCGCGGTT 20810  
 QY 181 lValIleGlnIleGlyLeuAla---LeuAlaLysAlaIleGlnThrLeuGlyAlaAla 200  
 DB 20809 ACATTTGCTCATTTCCGAAGGCAATCTGAACCAATCAAAACCC-----A 20759  
 QY 200 hLysSerAlaLeuSerSerThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysL 220  
 DB 20758 CCAACGACCCCGAGTATGCTTATTGAAACAACCTTCAAGTTGCAAAAACATCAAC--- 20703  
 QY 220 euGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgLysGlnGlnMetL 240  
 DB 20702 -----TGAATTCAGGTGACGCTTGTACGATAAAGGAGCTCAACAG----- 20658  
 QY 240 yAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValM 260  
 DB 20657 -----GAGCGCATGACACCCCGACGACGACGCTGCG 20627  
 QY 260 eTlleAlaValSerValAlaIleThr-----ValIleSerIleValAlaAlaIlePhe 278  
 DB 20626 TCGTTATCTGCTTAACCGATTGACCTTACGGCGCATCTCCCGCGGACGCGCGGAA 20567  
 QY 278 hrcCGSLyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGlyAlaAlaAlaG 298  
 DB 20566 CGCGGGCGCGCGGACGAGGA---GCGGAGAGACCGCCACAGAACGACGACGCGGAA 20510  
 QY 298 lYgLy---AlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 317  
 DB 20509 CTGAGATACGACAGCAAGCGACGACGACGACGAGTACGACGACGACATCAGCTCGAG 20450  
 QY 317 IaValValGlnAlaValLysGlnAlaValAlaIleThrAlaValAlaArgGlnAlaIleThrAla 337  
 DB 20449 CTATCACACACCGCCAGCGCAAGCGGCACTGCGCAGTCTC-----G 20408

PI Chen Y, Zychlinsky A:  
 XX WPI: 1997-393276/36.  
 DR P-PSDB: AAW29893.  
 XX  
 PT Shigella ipab protein induces apoptosis by binding to  
 PT interleukin-1-beta converting enzyme - useful for treating cancer,  
 PT auto-immunity, inflammation, etc. by eradication of unwanted cells  
 PS  
 Claim 3: Page 67-69; 108pp: English.  
 XX  
 XX This sequence comprises DNA encoding the Ipab protein (see AAW29893)  
 CC of Shigella flexneri. This protein induces apoptosis by binding to  
 CC interleukin-1-beta converting enzyme (ICE). A claimed method of  
 CC inducing apoptosis in a eukaryotic cell, thereby killing the cell,  
 CC comprises: (a) providing to a cell to be killed a DNA molecule (1)  
 CC in expressible form which encodes the Shigella ipab protein or a  
 CC functional derivative; and (b) causing (1) to be expressed in the  
 CC cell, thereby producing the Ipab protein or derivative which  
 CC induces apoptosis and kills the cell. This approach is useful in  
 CC treating e.g. cancer, autoimmunity, inflammation and chronic viral  
 CC infections. Protein or peptide fragments, and the DNA encoding  
 CC them, can also be used to treat diseases that involve apoptotic  
 CC mechanisms in their pathogenesis, e.g. AIDS, degenerative diseases  
 CC such as Alzheimer's disease, myelodysplastic disorders, ischaemic  
 CC injury or toxin-induced liver disease. Methods of gene therapy  
 CC relying upon controlled expression of Ipab in a target cell are  
 CC disclosed. Also claimed are methods for: inhibiting the  
 CC interaction of an apoptosis-inducing protein or peptide with ICE;  
 CC detecting a compound capable of inhibiting the binding of Ipab to  
 CC ICE; screening a candidate protein for its ability to interact  
 CC with Ipab in a cell; and treatment of Shigellosis. Homologues of  
 CC Ipab such as yopB of Yersinia enterocolitica and Yersinia  
 CC pseudotuberculosis, and sipB of Salmonella typhimurium and  
 CC Salmonella typhi can also be used in the methods of the invention.  
 XX  
 SQ Sequence 1743 BP: 632 A: 360 C: 302 G: 449 T: 0 other:  
 Alignment Scores:  
 Pred. No.: 9,91e-05 Length: 1743  
 Score: 185.00 Matches: 105  
 Percent Similarity: 38.10% Conservative: 84  
 Best Local Similarity: 21.17% Mismatches: 171  
 Query Match: 7.81% Indels: 136  
 DB: 18 Gaps: 21  
 US-09-889-314-2 (1-496) x AAT85939 (1-1743)  
 QY 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMet----- 55  
 DB 316 AAGTCCACGACACGCAACGCAACGCAAAAAACCTAGAAATTCCTCGATAAATTAACACT 375  
 QY 55 ----- 55  
 DB 376 CTTCATCTGAACCTGAAGACTAACCCAGACTATGAAAAACAAATTAACTAATAA 435  
 QY 56 GluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLys 75  
 DB 436 AAGCGAGATTCTTAATAAAGACCTAGAAATAATTAACCAATTAACACAGATTA 495  
 QY 76 ThGlnGluThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLys 95  
 DB 496 TCGAACCTCGATCCA-----GAGTCACCCAGAAAAAGAAAAA 531  
 QY 96 AlaGlyAlaAspThrGlyValSerGlyAlaAlaIleThrThrAlaSerSerThrAlaThr 115  
 DB 532 TTAAGCCGGAGAA-----ATACAACTCAGCATCAAAAAAGCGAGAGTT 579  
 QY 116 LysIleAlaMetGlnThrSerIleGlnGluAlaSerLysSerMetGluSerThrLeuGln 135  
 DB 580 AAA-----GACAGGACATTGATTGACGCAAAACCTGCAATTCATAGC----- 624  
 QY 136 SerLeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAla 155

DB 625 AAACCTTACAGTAATATCATGCACTCGCAAAAAAGAAATAGACTCT-----TTTTCGCA 678  
 QY 156 LeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGly 175  
 DB 679 TTTTCAAACACACCAATCT---GCTGACAGCTATCCACCAGCAAGCAAAATCAATTAACCGCA 735  
 QY 176 ValThrProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195  
 DB 736 CTTCGACGTTACTCAATG-----ATGGCACTTTATTCACACTA 777  
 QY 196 LeuGlyAlaIleThrLysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAsp 215  
 DB 778 GTTGGAAAAAATATGCAAAACATTTAAATAATGATGCTGCTATTCAGCTCTCCAA 837  
 QY 216 GlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGln 235  
 DB 838 GAATCAAGAAAAACCTGAATGTGAGAAAAATCT-----GATAG 876  
 QY 236 TyrGlnGluMetLysAlaAlaGlu---GlnLysSerLysAspLeuGlnGlyThrMetAsp 254  
 DB 877 TAT-----GCTGCTGAGTACTAAAGCAAGAACTCAGAGATATATGGGT 924  
 QY 255 ThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAla 274  
 DB 925 TGTGTTGGGAAATATCTGGGCACTT-----TTTACTATGTTATGTTGTGCA 975  
 QY 275 AlaIlePheThrGlyGlyAlaGlyLeuAlaGlyLeuAlaGlyValAlaValAlaGlyAla 294  
 DB 976 GCAGCTTTTCTGAGAGAGCTCTGTACAGCTGCGAGCTGTGGTTACTCTT----- 1029  
 QY 295 AlaAlaAlaGlyLysAlaAlaGlyAlaAlaAlaIleThrThrValAlaThrGlnIleThr 314  
 DB 1030 -----ATGGTT 1035  
 QY 315 ValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIle 334  
 DB 1036 ACGGATCTCTATGATACAGCAAGCAGCGCGC----- 1065  
 QY 335 ThrAlaIleIleLysAlaIleValLysSerGlyIleLysAlaPheIleLysThrLeuVal 354  
 DB 1066 AATTCCTTCATGACAAACGCCCTGATCCGATCATGAAACAGCTCATTTGAACCTTAATC 1125  
 QY 355 LysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGln 374  
 DB 1126 AAATCTCTTCAGATCATTTTACAAA-----ATGCTGGAAGGCTTGGGC 1170  
 QY 375 MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrp 394  
 DB 1171 GTGCACTCGAAAAAGCCAAATATGATTGGCTTCTTCTGGGGGCAATCCGAGCGGCTT 1230  
 QY 395 ValThrValGlyValGlyValAlaValAlaIleProAlaLeuGlyLysGlyIleMetGln 414  
 DB 1231 GTCCTAGTTGACAGAGCTGCTCTCGTAGCC-----ACTGTGTGTA 1272  
 QY 415 MetGlnLeuSerGluMetGlnGlnAsnValAlaGlnIleGlnLysGlnValGlyLysLeu 434  
 DB 1273 -----CAGGACACGCAAAAACCTGCGAGAAAATATTGGCAATA 1311  
 QY 435 -----GlnAlaAlaAlaAspMetIleSerMetPheThrGln---PheTrpGlnAla 451  
 DB 1312 ATAGGTAACCCCTCACAGACTTATACCAAGTTTCTTCAGAMTTTCTTCCACTG 1371  
 QY 452 SerLysIleAlaSerLysGlnThrGlyLysSerAsnGluMet----- 465  
 DB 1372 GACGATTTAATCACAATGCTGTTGCCAGATTAAATTAATTAATTTCTGTGCGCGGTGAT 1431  
 QY 466 -----ThrGlnLysAlaThrLysLeuGlyAla 474  
 DB 1432 GAAGTAATATCCAAACAAATTTTCCACCACATTTAAACCAAGCAGTTTATTGAGGAA 1491  
 QY 475 GlnIleLeuValAlaTyrAlaAlaIleSerGlyAlaIleAlaGlyAla 490

Percent Similarity: 37.00% Conservative: 76  
 Best Local Similarity: 23.29% Mismatches: 223  
 Query Match: 7.85% Indels: 126  
 DB: 18 Gaps: 25

US-09-889-314-2 (1-496) x AAT85043 (1-7766)

```

OY 5 SerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr 24
   :::::||||| ::::: ||| ::::: ||| ::::: |||
Db 52 GCTGTAGTGCCTGTAAGTGG-----GACACATATTAAAGTCGATTGTTAAG 99
OY 25 SerThrProGlnGlyValProGlnLysPylsSerGlyAsnGlnLysGlnIle 44
   ::::: ||| ::::: ||| ::::: |||
Db 100 GCTGCT-----GCTGAGGCTGCCAGGATGARAAGACCTGGGAGAGCTAAATCCG 153
OY 45 GlnGlnThrArgGlnGlyLysAsnThr-----GlnMetGlu 56
   ::::: ||| ::::: |||
Db 154 ATTGCTGCTGCTATTGGGAGGTAATGAGATGGTGGAGCTTAAGCATGATGAAG 213
OY 57 SerAspAlaThrIleAlaGlyAla-----SerGlyLysAspLysThrSer 71
   ||| ||| ||| |||
Db 214 AAGGATGATCAGATTGCTGCTCTATTGCTTGAGGGGAGCTAAGATGGAAGTTT 273
OY 72 SerThrThrLysThrGlnThrAlaProGlnGlnGlyVal-----AlaAlaGlySerGlu 89
   ::::: ||| ::::: |||
Db 274 GCTGTAAGAAATGATGAGAAAGGAGGCTGAGGGGCTATTAAAGGAGCTGGCAGTTG 333
OY 90 SerSerGlnSerGlnLysAla-----GlyAlaAspThrGlnLysSerGly 104
   ::::: ||| ::::: |||
Db 334 TTGGATTAAGCTGTAAAGAGCTTAAAGACAGCTGAGGGGCTTCAAGTCTGCTGCA 393
OY 105 AlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlu 124
   ||| ||| ||| |||
Db 394 ATTGCAAGAACTTGCTGCTGCTGCTATTGCTGCAAGGTTGCTGATGAAGCCGAGGAAG 453
OY 125 GlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeu-----GlnSerLeuSer 141
   ::::: ||| ::::: |||
Db 454 GGGATTGCTTAAGCGGATTAAGAGAGATTGTTGAAGCTGCTGGGGGAGTAAAGAGCTGAAA 513
OY 142 AlaAlaGlnMetLysGln-----ValGlnAlaValValAla 154
   ||| ||| ||| |||
Db 514 GTTGTGCTGCTTAAGAGGCAATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 573
OY 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnLysProGlnLeuProLysPro 174
   ||| ||| ||| |||
Db 574 GCTCATGCTGGGGACAGTACAGCTGCTACCAAG----- 606
OY 175 GlyValThrProArgSerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln 194
   ||| ||| ||| |||
Db 607 -----GCGGCTGCTGCTGTTAGT 624
OY 195 ThrLeu-----GlyGlnAlaThrLysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGln 213
   ::::: ||| ||| ||| |||
Db 625 GCTGTTAGTGGGACAGATATTAAAGTGGATTGAAGGCTGCTGGCTGGCTGGCTGGT 684
OY 214 AlaAspGlnThrAsnLysLeuGlyLeuGlnLysGln-----AlaIleLysIleAspLys 231
   ::::: ||| ||| ||| |||
Db 685 GATCAGAGGAGAAAGAGGCTGGGATGCTAATAAATCCATTCCTCTCTATTGGGAAG 744
OY 232 ---GlnArgGlnGlyLysGlnLysLeuLysAlaAlaGlnLysSerLysAspLeuGln 250
   ::::: ||| ||| |||
Db 745 GGTGATGCGAGAAATGTCGCGAGTTTAAATCATGATGAGATGAAGAAGATGAT----- 798
OY 251 GlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIle 270
   ||| ||| ||| |||
Db 799 -----CAGATTGCTGCTGCTATTGCTTTGAGGGGAGT 831
OY 271 SerIleValAlaAlaIle-----PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly 289
   ::::: ||| ||| ||| |||
Db 832 GCTAAGAGATGAAGATTCTCTGAAGAGTGGTGGTGGAGAAAGGAGGAGGAGGAGG 891
OY 290 AlaAlaValGlyAlaAla-----AlaAlaGlyLysAlaAla 301
   ||| ||| ||| |||

```

```

Db 892 CCTATTAAAGGAGCTGCTGAGTTGTGATTAAGCTGTAAAGCTGTAAAGACAGCTGAG 951
OY 302 GlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAla---ValValGln 320
   ||| ||| ||| ||| |||
Db 952 GGGGCTTCAAGTGTATGATGCAATTGGAGAAATTGTCGCTAATGCTGCTGCTGCAAG 1011
OY 321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
   ::::: ||| ||| ||| |||
Db 1012 GTTGTGATTAAGCCAGCTGACGCGGATTCCTAAGGAGATTAAGAGATTTGTAAGCT 1071
OY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
   ||| ||| ||| |||
Db 1072 GCTGGGGGAGTGAA-----AAGCTGAAGATTGCTGCTACAGGGGAG 1116
OY 361 IleSerLysGlyIleSerLysValPheAlaLys---GlyThrGlnMetIleAlaLysAsn 379
   ::::: ||| ||| ||| |||
Db 1117 AGTAATTAAGGGGCAAGGAGATTGTTGGAGGCTGCTGCTGCTAATGCTGGGAGC 1176
OY 380 PheProLysLeuSerLys-----ValIleSerSerLeuThrSerLysTrpValThr 396
   ||| ||| ||| |||
Db 1177 AGTAGAGCTGCTAGCAAGGCGGCTGCTGCTTATGCTGCTTATGCTGAGGAGCATATTA 1236
OY 397 ValGlyValGlyValValAla----- 404
   ::::: ||| ||| |||
Db 1237 AGTCGATTT---GTTAAGGCTGCTGATGCGCTGATCAGAGAGGAAAGACCTGGGAGT 1293
OY 405 -----AlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlu 419
   ||| ||| ||| |||
Db 1294 GCTACAAATCCGATTGCTGCTCTATTGGGAAGGT----- 1329
OY 420 MetGlnAlaAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAsp 439
   ::::: ||| ||| ||| |||
Db 1330 AATGAGAAATGTCGCGAGTTTAAGATGAGATGAAGAGATGATGATGCTGCT 1389
OY 440 MetIleSerMetPheThrGlnPheTrpGlnAlaSerLysIleAlaSerLysGlnThr 459
   ||| ||| ||| |||
Db 1390 GCTATTGCTTTG-----AGGGGAGATGCTTAAGATGAAGATTGCTGTAAGAGATGCT 1443
OY 460 GlyGlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeu----- 477
   ||| ||| ||| |||
Db 1444 GGTGAGAAAGGAGAG---GCTGAGGGGCTATTAAAGGAGCTGCTGATGTTGGATTAAG 1500
OY 478 -----LysAlaTyrAlaAlaIleSerGlyAlaIleAlaGly 489
   ||| ||| ||| |||
Db 1501 CTGTTAAAGCTGTAAAGACACTGAGGGGCTTCAAGTGT 1542

RESULT 25
AAT85939 standard; DNA; 1743 BP.
AC AAT85939;
DT 10-MAR-1998 (first entry)
DE Shigella flexneri ipav gene.
KW IPav: apoptosis; interleukin-1-beta converting enzyme; ICE; cancer;
KW autoimmune disease; inflammation; AIDS; degenerative disease;
KW Alzheimer's disease; myelodysplastic disorder; ischemic injury;
KW toxin-induced liver disease; gene therapy; shigellosis; ss.
OS Shigella flexneri.
PN WO9726790-A1.
PD 31-JUL-1997.
PF 10-JAN-1997; 97WO-US00399.
PR 25-JAN-1996; 96US-0591079.
XX (UUNY ) UNIV NEW YORK STATE.
XX
XX

```



```

OY 175 -----GlyValThrProArgSerGluValIleGluIle 185
Db 7696 GCGAACAAGATGTTGCCAAGGAGCTGCACGCCCA--GCAGCAGCAGCAGCAGCA 7640
OY 186 GYLLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyLysAlaThrLysSerAlaLeuSer 205
Db 7639 GCAGCAGCCCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 7583
OY 206 AsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlyLys 224
Db 7582 GAGAGCAATGTGACCCCAATCATCATCATCATCATCATCATCATCATCATCATCATCAT 7523
OY 225 -----GlnAlaIleLysIleAspLysGluArgGluGlu 235
Db 7522 GCCAACAATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 7463
OY 236 -----TyrGlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeuGlyThrMet 253
Db 7462 CAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7403
OY 254 AspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 273
Db 7402 ACCAACAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7346
OY 274 AlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValGly 293
Db 7345 GCTGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7286
OY 294 AlaAlaAlaAlaGlyAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 313
Db 7285 TCGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7226
OY 314 ThrValGlnAlaValAlaGlnAlaValLys 323
Db 7225 ACAGTAATGCAATTTGGCACAGTGTCTGCGCA 7196

RESULT 24
AAT85043
ID AAT85043 standard; DNA; 7766 BP.
AC AAT85043;
XX
XX
XX 22-FEB-1998 (first entry)
XX
DE Borrelia variable major protein (VMP)-like sequence (vls) locus.
XX Variable major protein-like sequence; VMP-like sequence; vls locus;
XX vlsE gene; Lyme disease; relapsing fever; therapy; diagnosis;
XX vaccine; ss.
XX
XX Borrelia burgdorferi strain B31-5A3 (ATCC 35210).
XX
XX Key Location/Qualifiers
XX 205..711
XX misc_feature
XX /tag= a
XX /label= vls2
XX /note= "truncated at 5' end"
XX 712..1293
XX misc_feature
XX /tag= b
XX /label= vls3
XX 1294..1869
XX misc_feature
XX /tag= c
XX /label= vls3
XX 1870..2439
XX misc_feature
XX /tag= d
XX /label= vls5
XX 2440..3009
XX misc_feature
XX /tag= e
XX /label= vls6
XX 3010..3483
XX misc_feature
XX /tag= f
XX /label= vls7
XX 3484..3990
XX misc_feature

```

```

FT FT /tag= g
FT FT /label= vls8
FT FT misc_feature
FT FT /tag= h
FT FT /label= vls9
FT FT 4549..5058
FT FT /tag= i
FT FT /label= vls10
FT FT 5059..5652
FT FT /tag= j
FT FT /label= vls11
FT FT 5653..6219
FT FT /tag= k
FT FT /label= vls12
FT FT 6220..6789
FT FT /tag= l
FT FT /label= vls13
FT FT 6846..7373
FT FT /tag= m
FT FT /label= vls14
FT FT 7274..7946
FT FT misc_feature
FT FT /tag= n
FT FT /label= vls15
FT FT /note= "the vls sequence provided is only
FT FT 7766 bases"
FT FT /tag= o
FT FT /label= vls16
FT FT /note= "the vls sequence provided is only
FT FT 7766 bases"
FT FT W09731123-A1.
FT FT
FT FT 28-AUG-1997.
FT FT
FT FT 20-FEB-1997; 97MO-US02952.
FT FT
FT FT 21-FEB-1996; 96DS-0012028.
FT FT
FT FT (TEXA ) UNIV TEXAS SYSTEM.
FT FT
FT FT Barbour AG, Hardham JM, Howell JK, Norris SJ, Weinstein GM;
FT FT Zhang J;
FT FT WPI; 1997-435172/40.
FT FT
FT FT Nucleic acid encoding variable major protein-like peptide of
FT FT Borrelia - useful for recombinant production of VMP like protein or
FT FT peptide, or for diagnosis of Lyme disease
FT FT
FT FT Claim 26; Page 101-105; 130pp; English.
FT FT
FT FT This DNA sequence comprises the variable major protein (VMP)-like
FT FT sequence (vls) locus of Borrelia burgdorferi. An infectivity
FT FT related 28-kb linear plasmid, pBB28la, of B. burgdorferi B31 was
FT FT isolated by subtractive hybridisation. It contained the vls locus,
FT FT which consists of 15 silent vls cassettes (vls2-vls16) and an
FT FT expressed vlsE gene (see AAT85043). Portions of several of the 15
FT FT silent vls cassette sequences, located approx. 500 bp upstream of
FT FT vlsE, recombine into the central vlsE cassette region during
FT FT infection, resulting in antigenic variation of the expressed
FT FT lipoprotein (see AAT22676) and hence immune evasion. Long-term
FT FT survival and pathogenesis in the mammalian host. Nucleic acids,
FT FT including vls2-vls16 sequences, encoding VMP-like proteins, can be
FT FT used for the recombinant production of VMP-like proteins, or for
FT FT the diagnosis of Lyme disease (claimed), and may also have
FT FT application in DNA vaccines.
FT FT
FT FT Sequence 7766 BP; 2029 A; 867 C; 3011 G; 1858 T; 1 other;
FT FT
FT FT Alignment Scores:
FT FT pred. No.: 0.00049 Length: 7766
FT FT Score: 186.00 Matches: 129

```



```

OY 131 GluSerThrLeuGluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAla 150
    :::: :::: |||| ::::: ::::: ::::
Db 2552 CAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2493
OY 151 ValValAlaAlaAlaLeu-----Ser 157
    :::: |||||
Db 2492 ATCGACACGACGACCAATGGTGGCCACAAACGAAACGAAATTCAGCGAAGCC 2433
OY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysPro----- 174
    :::: ::::: ::::: ::::: ::::: :::::
Db 2432 GCTGAGCCACGAGGCAAGTCAAAAGGTGCAATCCGCAATACGACAAACGAAAGTCCC 2373
OY 175 -----GlyValThrProArgSerGluValIleGluIle 185
    :::: |||||
Db 2372 GCGCAAAACAATGTTGGCAAAAAGAAAGCTGCACGCCA---GCGACGACGACGACGAC 2316
OY 186 GlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThrLysSerAlaLeuSer 205
    :::: |||||
Db 2315 GCAGCAGCGCGCGCGCACGACGACGACGACGACGACGACGACGACGACGACGACG 2259
OY 206 AsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLys--- 224
    :::: |||||
Db 2258 GAGAGCAATGTGACCCCAATCAATGCGATTTATTAACGCTGCCGCCCAAGAGCG 2199
OY 225 -----GlnAlaIleLysIleAspLysGluArgGluGlu--- 235
    :::: |||||
Db 2198 GCCAAACAATCACCACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2139
OY 236 -----TyrGlnGluMetLysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMet 253
    :::: |||||
Db 2138 CAGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2079
OY 254 AspThrValaSerThrValaMetIleAlaValaSerValaIleThrValIleSerIleVal 273
    :::: |||||
Db 2078 ACCACAACGCGCCACAAACA---ATAACAATGTGTGCTGCTGTAAATGCTGCTGCT 2022
OY 274 AlaAlaIlePheThrCysGlyValaGlyLeuAlaGlyLeuAlaAlaGlyValaValaGly 293
    :::: |||||
Db 2021 GCTGTGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1962
OY 294 AlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAla 313
    :::: |||||
Db 1961 TCGCAGCAGCAGCAACAGACGCGCGGCAAGCGCGCAATTAATTAACAAGTTGCA 1902
OY 314 ThrValGlnAlaValaValaGlnAlaValaLys 323
    :::: |||||
Db 1901 ACAGTATCGAATTTGGCACAGTGTGCGA 1872

RESULT 23
ABL04818/C
ID ABL04818 standard; cDNA; 18869 BP.
XX
XX ABL04818;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 8936.
DE
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
KW
XX
XX pharmaceutical; gene; ss.
OS
XX
XX Drosophila melanogaster.
PM
XX
XX WO200171042-A2.
PD
XX
XX 27-SEP-2001.
PF
XX
XX 23-MAR-2001; 2001WO-US09231.
PR
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX
XX 11-JUL-2000; 2000US-0614150.
XX

```

```

PA (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW.
PI
XX WPI: 2001-656860/75.
DR
XX P-PSDB: ABB60715.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 8936; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB5737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 18869 BP; 5494 A; 3649 C; 3579 G; 6147 T; 0 other;

Alignment Scores:
Pred. No.: 0.00129 Length: 18869
Score: 186.50 Matches: 80
Percent Similarity: 38.00% Conservative: 53
Best Local Similarity: 22.86% Mismatches: 162
Query Match: 7.87% Indels: 55
DB: 23 Gaps: 10

US-09-889-314-2 (1-496) x ABL04818 (1-18869)
OY 25 SerThrProGlnGlyValProGlnGlnAspLysSerGlyAsnGluThrLys----- 42
    :::: |||||
Db 8233 ACACACCCACCAAAACACCCACCAAGTT---AAGCAGACGACGACGACGACGACCA 8177
OY 43 -----GlnIleGlnGlnThrArgGlnGlyLysAsnThr 53
    :::: |||||
Db 8176 CCAGCAGCACCACCAACCAATGCGCAAGATCAGCATGACCAACCAATGCGACGACCACT 8117
OY 54 GluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerThr 73
    :::: |||||
Db 8116 GTTGCCGAGGCGGAGAAAACAATCTCACCGAAGGGGTTGMAACAATCTGCGCAACCCGTA 8057
OY 74 -----ThrLysThrGluThrAlaPThrGlyValSerGlyAlaAlaAlaThrAla 110
    :::: |||||
Db 8056 GCCATGCCACGCGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 7997
OY 91 SerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThrAla 110
    :::: |||||
Db 7996 AATTGACGAGAAACGACGAGCCCAAAATTAATACGCGGACCGCCCAAGTCACCTGAA 7937
OY 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMet 130
    :::: |||||
Db 7936 CAACAACGCGCAACGCGCACCCAGTCCGAAGAAATTAATTAACAACCACTGACCAAGAA 7877
OY 131 GluSerThrLeuGlnSerLeuSerLeuSerAlaAlaGlnMetLysGluValGlnAla 150
    :::: |||||
Db 7876 CAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 7817
OY 151 ValValAlaAlaAlaLeu-----Ser 157
    :::: |||||
Db 7816 ATCGACACGCGACCAATGTTGGCCACAAACGAAACGAAATTCAGCGAAGCC 7757
OY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysPro----- 174
    :::: |||||
Db 7756 GCTGAGCCACGAGGCAAGTCAAAAGGTGCAATCAGCAATATGACGACAAAGAGTCCC 7697

```

```

Db 2980 GTTCCGAGCGAGAAACATCTCACGAGGGGTGAACCAATCTGGACACGCCGA 3039
QY 74 -----ThrlsThrlglnThrlaProGlnGlnGlyValAlaAlaGlySer 90
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3040 GCCATGCCAACGCCACACCAACCAACCAACCAACCAACCAACCAACCAATCA 3099
QY 91 SerGlnSerGlnLysAlaGlyAlaSerThrGlnValSerGlyAlaAlaThrThra 110
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3100 AATTGAACGAGCAACGAGGCGCAAAATTAATAGCGGCGCGCGCAAGTCACTGA 3159
QY 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnLysLysSerMet 130
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3160 CAACAACAGGCAACGAGCGCGCGCAACCAAAATTAATTAACCACTGACAAAGAA 3219
QY 131 GluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluVal 150
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3220 CAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3279
QY 151 ValValValAlaAlaAlaLeu-----Ser 157
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3280 ATGCACACGCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3339
QY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro----- 174
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3340 GCTGAGCCCAACGCGCAAGTCAAAAGTGCATACCGCAATACGACAAAGCAAAAGTCC 3399
QY 175 -----GlyValThrProArgSerGluValIleGluIle 185
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3400 GGCAAAACAAATGTTGGCAAAACAAAGCTGCACGCCA---GCAGCAGCAGCAGCAGCA 3456
QY 186 GlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysLysLysSerAlaLeuSer 205
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3457 GCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3513
QY 206 AsnThrAlaSerThrGlnAlaGlnAlaSerGlnThrAsnLysLysLysLysLys 224
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3514 GAGAGCAATGTGACCCCAATCATCTGATCAATTTATACGCTGGCCCCCAACAAGCG 3573
QY 225 -----GlnAlaIleLysIleAspLysGluArgGluGlu--- 235
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3574 GCCAAACATCATCCGACAGCAGCAATGTCATTCGCAATTCGCAACAGCAGCAGCAG 3633
QY 236 -----TyrGlnGlnMetLysAlaIleGlnLysSerLysAspLeuGlnGlyThrMet 253
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3634 CAGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3693
QY 254 AspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 273
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3694 ACCACACGCCGCCACACA---ATAACAAATGTGTGTGTGTGTGTGTGTGTGTGT 3750
QY 274 AlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleAlaIleValaGly 293
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3751 GCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3810
QY 294 AlaAlaIleAlaGlyGlyAlaIleAlaGlyAlaIleAlaIleAlaIleThrValaIleThrGlnIle 313
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3811 TCGCTCAGCAGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3870
QY 314 ThrValGlnAlaValValGlnAlaValLys 323
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX 3871 ACAGTATCGAATTTGGCAGCTGTGTGCGA 3900

```

```

KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceuticall; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 4033; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
XX (ABBS7737-ABBS7072)
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6771 BP; 1504 A; 1470 C; 1519 G; 2278 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.000382 Length: 6771
XX Score: 186.50 Matches: 80
XX Percent Similarity: 38.00% Conservative: 53
XX Best Local Similarity: 22.86% Mismatches: 162
XX Query Match: 7.87% Indels: 55
XX DB: 23 Gaps: 10
XX
XX US-09-889-314-2 (1-496) x ABL17520 (1-6771)
XX
XX QY 25 SerThrProGlnGlyValProGlnAspLysLeuSerGlnLysLys----- 42
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX Db 2909 ACAACACCAACAAACACCAACGTT---AAGCACAACAGATGATGATGATGATGATCA 2853
XX
XX QY 43 -----GlnIleGlnGlnThrArgGlnGlnLysAsnThr 53
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX Db 2852 CCAGCAGCAGCAACGACCAATGCGCAAGTCAAGCATGACACAAAGCGCAGCAACT 2793
XX
XX QY 54 GluMetGlnSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThr 73
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX Db 2792 GTTGGCGAGCGAGCAACCAATCATCCAGCAAGGGGTGAACAATGTGGCAAGCGCGTA 2733
XX
XX QY 74 -----ThrlsThrlglnThrlaProGlnGlnGlyValAlaAlaGlySer 90
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX Db 2732 GCCATGCCAACGCCACCAACCAACCAACCAACCAACCAACCAACCAATCA 2673
XX
XX QY 91 SerGlnSerGlnLysAlaGlyAlaSerThrGlnValSerGlyAlaAlaIleThrThra 110
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX Db 2672 AATTGAACGAGCAACGAGGCGCAAAATTAATAGCGGCGCGCGCAAGTCACTGAA 2613
XX
XX QY 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnLysLysSerMet 130
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 4033.
XX
XX

```





```

XX MPI: 2000-679515/66.
DR P-PSDB: AAB19772.
XX
PT Detecting small integrin-binding ligand N-linked glycoproteins for
PT detection of a tumor or protection against a complement mediated immune
PT response, comprises detection where Factor H is not an inhibitor -
XX
PS Disclosure: Page 98-100; 110pp; English.
XX
CC The present sequence is that of DNA encoding human dentin
CC sialoprophosphoprotein (DSPP, see AAB19772), a member of the small
CC integrin binding ligand, N-linked glycoproteins (SIBLINGS) family.
CC The invention provides methods and compositions for exploiting the
CC discovery that members of the SIBLINGS family bind to complement
CC Factor H, conferring resistance to complement mediated lysis. A
CC claimed method of conferring protection against a complement
CC mediated immune response involves providing a reservoir or other
CC supply in the subject's body so that a SIBLINGS protein can be
CC dispersed to interfere with complement mediated lysis and
CC inflammation. This protects cells that are grafted onto foreign
CC tissue or bone marrow cells introduced into a foreign host. The
CC SIBLINGS protein can be BSP, OPN, DMP1 or DSPP. A method of
CC detecting a SIBLINGS protein in a sample from a subject suspected
CC of having abnormal bone turnover, especially osteoporosis, is also
CC claimed.
XX
SQ Sequence 8201 BP; 2801 A; 1477 C; 1873 G; 2050 T; 0 other:
XX
Alignment Scores:
Pred. No.: 0.0004 Length: 8201
Score: 187.00 Matches: 106
Percent Similarity: 39.24% Conservative: 91
Best Local Similarity: 21.12% Mismatches: 251
Query Match: 7.89% Indels: 56
DB: 21 Gaps: 11
XX
US-09-889-314-2 (1-496) x AAA88864 (1-8201)
OY 1 AASPThASmMetSerIleSerSerSerGlyProASPasnGlnLysAsnIleMetSer 20
DB 6394 GACAGTAGGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6453
OY 21 GlnValIleThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
DB 6454 AGTGATAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6513
OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
DB 6514 AGCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6573
OY 61 IleAla-----GlnAlaSerGlyLysAspLysThr 70
DB 6574 GACAGCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6633
OY 71 SerSerThrThrLysThrGlnThrAlaProGlnGlnGlyValAlaIleGlnLysGluSer 90
DB 6634 GATAGCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6693
OY 91 SerGluSerGlnLysAlaGlyAlaAspThrGly-ValSerGlyAlaAlaAlaThrThrAl 110
DB 6694 GACAGCAGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6753
OY 110 AsSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMet 130
DB 6754 AGCAACAGCAGTAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6813
OY 130 tGluSerThrLeuGlnLysLeuGlnSerLeuSerAlaIleGlnMetLysGlnValGluAl 150
DB 6814 AGCAGTAGACA-----GCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6858
OY 150 aValValValAlaIle--AlaLeuSerGlyLysSerSerGlySerAlaLysLeu----- 167

```

```

DB 6859 AGTGATAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6918
OY 168 -----ThrProGlnLeuProLysProGlnGlyValThrProArgSerGlnVal 182
DB 6919 ACCAGTAATAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6978
OY 182 ILeGlnIleGlyLeuAlaIleAlaLysAlaIleGlnThrLeuGlnGlyGluAlaThrLysSe 202
DB 6979 GATAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7037
OY 202 rAlaLeuSerAsnTyrrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLe 222
DB 7038 CAGTAGACAGCAGTAGATAGCAATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7097
OY 222 uGlnLysGlnAlaIleLysIleAspLysGlnLarGlnGluGlyTrpGlnGlnMetLysAlaAl 242
DB 7098 CAGTAGACAGCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7157
OY 242 aGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThr-ValAsn----- 257
DB 7158 TAGCAGTAGACAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7217
OY 258 --ThrValMetIleAlaValaIleSerValaIleThrValIleSerIleValaIleAlaIle- 276
DB 7218 TGACAGCAGCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7277
OY 277 -----PheThrGlyGlnAlaGlyLeuAlaGlyLeuAlaGlyAlaIleAlaValaIleG 293
DB 7278 CAGTGACAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7337
OY 293 lYAlaAlaIleAlaIleGlyGlyAlaIleAlaIleAlaAlaAlaAlaThrThrValaIleThrGlnI 313
DB 7338 CAGCAGTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7397
OY 313 lThrValaIleAlaValaIleAlaValaIleGlnAlaValaIleThrAlaValaIleArgGlnA 333
DB 7398 TGACA-----GCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7448
OY 333 lAlleThrAlaIleAlaIleLysAlaIleAlaValaLysSerGlyIleLysAlaPheIleLysThrL 353
DB 7449 CAGTGACAGCAGCG-----ACAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7502
OY 353 euValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyLT 373
DB 7503 CAGCGATAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7544
OY 373 hGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLT 393
DB 7545 CAGCAGTAGATAGCA-----GTGACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 7577
OY 393 yStrpValThrValaIleGlyValaIleGlyValaIleAlaIleAlaIleAlaLeuGlyLysGlyIleM 413
DB 7578 CGATAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7637
OY 413 eGlnMetGlnLeuSerGlnMetGlnGlnGlnAsnValaIleAlaGlnPheGlnLysGlnValaIleG 433
DB 7638 CAGTGACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7697
OY 433 ySLeuGlnAlaIleAlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaIleSerLT 453
DB 7698 CAGCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7756
OY 453 ySleAlaSerLysGlnThrGlyGlnSerAsnGlnLumetThrGlnLysAlaThrLysLeuG 473
DB 7757 GGTATAGCAGTAGACAGACATCTGACAGCAATGATGAGTAGACAGCAGAGAGAGAGAG 7816
OY 473 lY 473
DB 7817 GT 7818

```

RESULT 20  
 ABO73537  
 ID ABO73537 standard; DNA; 8201 BP.



```
Dh 10056 TAGACGCCAAAAAGCCGAGATT-----TCGCGCATCAGCTACAGCAAGCAC 10106
Oy 67 ysaSplThSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAla 87
Dh 10107 ACGACACCACCACTCATGTAAGAAC--GCCGCTCCCTCAAGGGTAGTGTGA---- 10159
Oy 87 LysGluSerSerGlnSerGlnLysAlaAlaAlaSerThrGlyAlaSerGlyAlaAla 107
Dh 10160 -----GANTCACCACCAACCCCAATGGGCTGGATACCAATGCAAGCCACACACT 10211
Oy 107 LathThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle---GluGua 126
Dh 10212 TTGAATCCACTGGGTGGCGCAACC-----ATACGCGCAGCGGTAGGTAGCAGG 10262
Oy 126 LaseLysSerMetGluSerThrLeuGluSerLeuGln---SerLeuSerAlaAlaGln 145
Dh 10263 CACGGGCGAGATGCCAAGATTATCTCGAAGGATCAAAAGCATTCACACAGAAACCG 10322
Oy 145 eLysGluValGluValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaL 165
Dh 10323 TGACGACAGCAAAATCTACTCTATGCAAAAACAGCAGCAGCGGCGACTAATCGAAA 10382
Oy 165 yslLeuGluThrProGluLeuProLysProGlyVal-----ThrProArgSerG 181
Dh 10383 CCTTGCATTTGCCGAGTTTCAACCGGCTCCCGCTGCTGCTGCGACCGCGCGGTT 10442
Oy 181 LuValIleGluIleGlyLeuAla---LeuAlaIleAlaIleGlnThrLeuGluAla 200
Dh 10443 ACATTGTGCACATTCGAAAGCAATCTGAAACCAATCAAAACCAACCCCTC-----A 10493
Oy 200 hrlYSerAlaLeuSerAsnTrpAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLys 220
Dh 10494 CCAAGCAGCCCGCATGCTCTATTGAAACAACATCAAGTTCGAAAAACATCAAC----- 10549
Oy 220 euGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGluTrpGlnGluMet 240
Dh 10550 -----TGGAATAGCGAGCTGCTTACGATTAATGAGCACTACAAACAG----- 10594
Oy 240 ysaAlaIleGluGlnLysSerLysAspLeuGluGlyThrMetLysPheValAlaThrVal 260
Dh 10595 -----GAGGCGATGACACCCGCGCAGCAGCAGCTGTCG 10625
Oy 260 eLlLeAlaValSerValAlaIleThr-----ValIleSerIleValAlaAlaIlePhe 278
Dh 10626 TCGTATGCTGCTATACCGTATGACGTACGCGTACGTGCTGCTGCTGCTGCTGCTG 10685
Oy 278 hrcYsGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaAlaAlaAla 298
Dh 10686 CGGCGGCGCGCGCAGCGCAGGA---CGGCGAGAGCGCGCAGCAGCAGCAGCGCGGA 10742
Oy 298 LysGly---AlaAlaGlyAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrVal 317
Dh 10743 CTGGAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 10802
Oy 317 laValValGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAla 337
Dh 10803 CTATACACACAGCCCGCAGCAAGCGCAGCTGCCAGCTC-----G 10844
Oy 337 laIleLysAlaAlaVal-----LysSerGlyIleLysAlaPheIleLysT 352
Dh 10845 CCAGCAGCCCGCTTCCCTCATCAACAAAGAGAGACATTAACCTATCCGTAAG 10904
Oy 352 hrlLeuValLys-----AlaIleAlaLysAlaIle 362
Dh 10905 AACTGGGCAAAAGCAGCACCGTCAAGACGGCGCACCGCCCTTAACCGCAGCGCTAC 10964
Oy 362 eLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleLysAsnIlePhe--- 380
Dh 10965 TGCAGGCGATTAAGCGGCTGAACACCAAGCAGCGCAAGCCCTCAGCAAACTTTTCAC 11024
Oy 381 --ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyVal 400
Dh 11025 GTCCCGCAGCAGCAAACTGACCGCTTAACCTGATCAACAGCACCGCTGCCGCAAGTTC 11084
```

```
Oy 400 LyValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlu 420
Dh 11085 ATACCGCATCAACGGCGCGCAGCTGAAAGACAACTGGGCGATGCCCACTGGGTGGA 11144
Oy 420 euGlnGlnAsnValAlaGlnPheGlnLysGluValGly-----LysLeuGlnAla 437
Dh 11145 TA-----GTGAGTACCGTACACGAGAGAAAGTACGACGAAATCAAAATTATATCA 11195
Oy 437 laAlaAspMetIleSerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSer 457
Dh 11196 GCGAAGACTACAT-----GCCACAGATAGCCCATG 11228
Oy 457 yslThrGlyLysSerAsnGluMetThrGlnLysAlaThr----- 470
Dh 11229 CCGTAGCAGGCTGTGCTGCGGCTAGCAATTAAGCAAAATGTCGGACGCGCAATGC 11288
Oy 471 -----LysL 472
Dh 11289 GCGCGCAGTGGCGCAGATGTTGGGAGAAACCTGTTGGACGAGCGATGTAGCACAAC 11348
Oy 472 euGlyAlaGlnIleLeuLysAlaTrpAlaAlaIleSerGlyAlaIleAlaGlyAla 490
Dh 11349 TGTCAACCCCAAGAACGCCAAAGATCATAGCTTACTCCAGATTATCGCAGCAGC 11404
RESULT 18
ID AAV52498 standard; DNA; 1436 BP.
XX AC AAV52498;
XX DT 23-OCT-1998 (first entry)
XX DE Streptococcus pneumoniae genome fragment SEQ ID NO:365.
XX KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
XX KW computer readable medium; vaccine; pharmaceutical composition; ds.
XX OS Streptococcus pneumoniae.
XX PN WO9818931-A2.
XX PD 07-MAY-1998.
XX PE 30-OCT-1997; 97WO-US19588.
XX PR 31-OCT-1996; 96US-0029960.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX PI Kunsch CA, Rosen CA;
XX DR WPI; 1998-272225/24.
XX PT Computer-readable medium with recorded Streptococcus pneumoniae
XX PT polynucleotide sequences - useful in diagnostic kits and assays, and
XX PT pharmaceutical compositions and vaccines for Streptococcus
XX PS pneumoniae
XX PS Claim 1; Page 1380-1381; 1409pp; English.
XX CC The present invention describes a computer readable medium which has
XX CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
XX CC recorded on it, or a representative fragment or a sequence at least 95%
XX CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
XX CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
XX CC Streptococcus pneumoniae. The present invention also describes an
XX CC isolated nucleic acid molecule encoding a homologue of any of the
XX CC fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the
XX CC nucleic acid molecule is produced by a process comprising: (a) screening
XX CC a genomic DNA library using as a probe a target sequence defined by any
XX CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
```

```

Db 1249 ----TCGCTTACAGCAGTACGCGCTCAGCAGCTTCAACAGCTTGGCTTCA 1305
OY 358 AAlaValAlaIleSerlySgIleSerlyValAlaIleLysGlyThrGlnMetIleAla 377
Db 1306 GCG-----TCAACGAGTGGCTGTGAGTCAAGCATCAACGAGTACGCTCAGCA 1356
OY 378 LysAsnProLysLeuSerlyValIleSerLeuThrSerlyStrpValThrVal 397
Db 1357 AGCAATCAGCTTGTGATCTGCATCAACAGCGCTCAGCTCAGCATCAGCAACGCGCC 1416
OY 398 GlyValGlyValValAlaAlaAlaProAlaLeuGlyLys-GlyIleMetGlnMetIle 417
Db 1417 TCAGCTTCAAGCAAGTACAGTGGCTCAGCGCTCAGCAAGTACCGCTTCAAGCTCAGCG 1476
OY 417 userGlnMetGln-----GlnAsnValAlaGlnPh 427
Db 1477 TCGACAAGTGGCTGGCTCAGCAGCTGATCTGAAATCGGCATCAACAGTGGCTCAGCT 1536
OY 427 eGlnLys----- 429
Db 1537 CACCAAGTACTAGTGCATCAGCTTACAGTCAACAGAGTGCATCGCTTGGCTCAACCA 1596
OY 430 -----GlyValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPh 447
Db 1597 GTGCATCAGAGTCAGCAAGTACAGTGGCTGCTACTTCC----- 1633
OY 447 eTrpGlnGlnAlaSerlySgIleAlaSerlySgIleThrGlyLysAsnGlnMetThrGln 467
Db 1634 ----GCATCAACAAGTCCCTGGCTTCAAGCAAGTCAAGCATCTCAATCAGCGTCAAC 1689
OY 467 nLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIle 487
Db 1690 CAGTGGCTTGGCTTCAAGCAAGTACAGTGGCTTCAAGTCAAGCAACAGCGCTCGGC 1749
OY 487 eAlaGlyAla 490
Db 1750 CTCAGCAAGC 1759

RESULT 17
AAV03553 ID AAV03553 standard; DNA: 15620 BP.
AC AAV03553:
XX 22-OCT-1998 (first entry)
DT DNA sequence that is specific for Neisseria meningitidis.
XX N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW meningitis; ss.
XX Neisseria meningitidis.
OS
XX
FH Key Location/Qualifiers
FT 1330..2973
FT /*tag= a
FT /*tag= "encodes AAW42633"
FT 3083..9028
FT /*tag= b
FT /*tag= "encodes AAW42634"
FT 9044..9475
FT /*tag= c
FT /*tag= "encodes AAW42635"
FT 10127..12121
FT /*tag= d
FT /*tag= "encodes AAW42636"
FT 12118..12609
FT /*tag= e
FT /*tag= "encodes AAW42637"
FT 12794..13066
FT /*tag= f
FT CDS

```

```

FT /*note= "encodes AAW42639"
FT 13297..14238
FT /*tag= g
FT /*note= "encodes AAW42640"
FT 14241..15176
FT /*tag= h
FT /*note= "encodes AAW42641"
FT W09802547-A2.
PD 22-JAN-1998.
XX 11-JUL-1997; 97WO-FR01295.
XX 12-JUL-1996; 96FR-0008768.
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (SMK ) SMITHKLINE BECHAM.
XX Nassif X, Tinsley C, Achtman M, Merker P, Ruelle J;
XX Vinals C;
XX WPI: 1998-110594/10.
XX P-PSDB: AAW42633, AAW42634, AAW42635, AAW42636, AAW42637, AAW42638,
XX AAW42639; P-PSDB: AAW42640, AAW42641.
XX Genes present in Neisseria meningitidis but not other Neisseria
XX species - and related host cells, RNA, anti-sense sequences,
XX polypeptide(s) and antibodies, useful for diagnosing Neisseria
XX meningitidis infection and in protective vaccines
XX Claim 8; Pages 67-78; 150pp; French.
XX AAV03518-53 represent sequences that are present in Neisseria
XX meningitidis but not in N. gonorrhoeae or N. lactamica, except for the
XX genes involved in biosynthesis of the capsule polysaccharide, trpA or C,
XX opc, porA, rotamase, sequence IC1106, Iga protease, pilin, pilQ,
XX proteins which bind transferrin and opacity proteins. The sequences are
XX found on chromosome 22491, mainly (or within 20 kb) between tufta and pilT
XX (region 1), pilQ and lambda-740 (region 2) or argF and opaB (region 3).
XX The DNA sequences are responsible for the differences in pathogenicity
XX between N. meningitidis and N. gonorrhoeae, specifically they include the
XX genes that allow N. meningitidis to cross the blood-brain barrier. DNA
XX sequences common to N. meningitidis and N. gonorrhoeae, but absent from
XX N. lactamica, are responsible for colonisation and penetration of the
XX mucosa. The DNA sequences can be used to produce probes and primers, and
XX antibodies produced against the encoded proteins are used in standard
XX hybridisation/immunoassay processes for diagnosis of N. meningitidis
XX infection, particularly meningitis.
XX Sequence 15620 BP; 4806 A; 3505 C; 3556 G; 3753 T; 0 other;
SQ

Alignment Scores:
Pred. NO.: 0 000437 length: 15620
Score: 191.50 Matches: 124
Percent Similarity: 38.33% Conservative: 83
Best Local Similarity: 22.96% Mismatches: 218
Query Match: 8.08% Indels: 115
DB: 19 Gaps: 24

US-09-889-314-2 (1-496) x AAV03553 (1-15620)
OY 17 AsnIleMetSerGlnValLeuThrSerThrProGlnGlyValProGlnGln----- 33
Db 9936 AATCTGGTGGCAGATCGACGCTACGCGCATTCATGATGCCCCCAAGGCGACATTA 9995
OY 34 ---AspLysLeuSerGlyAsn-GluThr-----LysGlnIleGlnGln 47
Db 9996 ACATCGAAGCCGGCGGCGGAATTTGACACTCTATGCGCTAGAGAGAGCTCACTACGACAAAC 10055
OY 47 rArgGlnGlyLysAsnThrGlnMetGlnUserAspAlaThr-IleAlaGlyAlaSerGlyL 67

```



0y	463	eserclyAlaIleAlagly- <u>Ala</u> HisLysThrAsnSn 495
		: : : :
Db	1648	GAGTGCCTCAGCTCAGCAAGCACATCAGCTTCTGAAT 1685
		: : : :
RESULT 16		
ID	AAV52497	
XX	AAV52497	standard; DNA; 2550 BP.
XX	AAV52497	
XX	AAV52497	
XX	23-OCT-1998	(first entry)
XX		
DE	Streptococcus pneumoniae	genome fragment SEQ ID NO:364.
XX		
KW	Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;	
XX	computer readable medium; vaccine; pharmaceutical composition; ds.	
XX		
OS	Streptococcus pneumoniae.	
XX		
PN	MO9818931-A2.	
XX		
PD	07-MAY-1998.	
XX		
PF	30-OCT-1997;	97WO-US19588.
XX		
PR	31-OCT-1996;	96US-0029960.
XX		
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;	
PI	Kunsch CA, Rosen CA;	
DR	WPI, 1998-272225/24.	
XX		
XX	Computer-readable medium with recorded Streptococcus pneumoniae	
PT	polynucleotide sequences - useful in diagnostic kits and assays, and	
PT	pharmaceutical compositions and vaccines for Streptococcus	
PT	pneumoniae	
XX		
PS	Claim 1; Page 1379-1380; 1409pp; English.	
XX		
CC	The present invention describes a computer readable medium which has	
CC	the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)	
CC	recorded on it, or a representative fragment or a sequence at least 95%	
CC	identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in	
CC	SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from	
CC	Streptococcus pneumoniae. The present invention also describes an	
CC	isolated nucleic acid molecule encoding a homologue of any of the	
CC	fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the	
CC	nucleic acid molecule is produced by a process comprising: (a) screening	
CC	a genomic DNA library using as a probe a target sequence defined by any	
CC	of the sequences in SEQ ID NO:1 to 391, identifying members of the	
CC	library which contain sequences that hybridise to the target sequence and	
CC	isolating the nucleic acid molecules from the members; or (b) isolating	
CC	mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid	
CC	molecules whose nucleotide sequence is homologous to amplification	
CC	primers derived from the fragment of the S. pneumoniae genome to prime	
CC	the amplification and isolating the amplified sequences. The computer	
CC	readable medium can be used in a computer-based system for identifying	
CC	fragments of the S. pneumoniae genome of commercial importance, or	
CC	expression modulating fragments of the S. pneumoniae genome. Products	
CC	from the present invention can be used in diagnosis kits and assays, and	
CC	pharmaceutical compositions and vaccines for S. pneumoniae.	
XX		
XX		
XX	Sequence 2550 BP; 630 A; 837 C; 584 G; 498 T; 1 other;	
Alignment Scores:		
Pred. No.:	4.69e-05	Length: 2550
Score:	192.00	Matches: 110
Percent Similarity:	33.97%	Conservative: 68
Best Local Similarity:	20.99%	Mismatches: 275
Query Match:	8.10%	Indels: 71

DB:	19	Gaps:	10
US-09-889-314-2 (1-496) x AAVS2497 (1-2550)			
Oy	5 SerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr	24	::: :::
Db	289 TCACACCAGTGGGTACAGCTTCAGATCGACAAGCGCCTCACGCTTCACGAAGAATCCAGTCT	348	::: :::
Oy	25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIle	44	::: :::
Db	349 TCACGCTCAGAGTCGACAAAGTGCGGCCCTCAACACAGCATGCATTGATCGGCATCAACC	408	::: :::
Oy	45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGlyVal	64	::: :::
Db	409 AGTCGCTCAGGCTTCAGCAACTGATGCTGATCAGCTTCAGCATCAGACAGTGCATCGCT	468	::: :::
Oy	65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal	84	::: :::
Db	469 TCACGATTCACCAAGTGGCTGCGCTTCACGCTCAACCAAGCGCTCACGTTCCAGCAAGTACC	528	::: :::
Oy	85 AlaAlaGlyLysGluSerSerGluLysAlaGlyAlaAspThrGlyValSerGly	104	::: :::
Db	529 AGTGCTTCAGCTCAGCATCAACAGAGTTCAGCTTCAGCATCAACAGAGTCCGCT	588	::: :::
Oy	105 AlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIleAlaMetGlnThrSerIle	123	::: :::
Db	589 TCACAGACACATCAGATGATGTAATCAGCGTCGACCAAGTGGCTGAGCTCAACCAAGTCCA	648	::: :::
Oy	124 ---GlnGlnAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAla	142	::: :::
Db	649 TCTGATTCGATCAACCAATGCGTTCAGCTTCAGCAAGTCAATGATCGATCGATCTCACCA	708	::: :::
Oy	143 AlaGlnMetLysGluValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGly	162	::: :::
Db	709 TCACAGAGTGCATGCGTTCGCGTCGACCAACAGTGCATCAGATCAGCAAGTACCGAGTCCG	768	::: :::
Oy	163 SerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProArgSerGluVal	182	::: :::
Db	769 TCAGCTTCGCCATCAACAACACT-----	789	::: :::
Oy	183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSer	202	::: :::
Db	790 -----GCCTCGGCTTCAGCAAGACATCAACATCGAATCAACGCTCAAC	834	::: :::
Oy	203 AlaLeuSerAsnTrpAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu	222	::: :::
Db	835 AGTGTCTCGGCTTCGACCAAGTACAGTTCGCTTCAGCTTCAGATCAAC-----AGCC	888	::: :::
Oy	223 GluLysGlnAlaIleLysIleAspLysGlnArgGlnGlyThrGlnMetLysAlaIle	242	::: :::
Db	889 TCGGCTTCAGCAACCACTCAGCTTCGATTCGCGCTCAACCAAGCGCTCGGCTCACCA	948	::: :::
Oy	243 GluGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla	262	::: :::
Db	949 AGCACTCAGCTTCGATGATGGGCTCAACCAAGCGCTCAAGCATCAACAGTAGTCT	1008	::: :::
Oy	263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyValActy	282	::: :::
Db	1009 TCGGCTTCAGCAACCAAGCGCTCGGTTGAGCATCAACAGTACGTCACCTTCACCG	1068	::: :::
Oy	283 LeuAlaGlyLeuAlaAlaGlyAlaValGlyAlaAlaIle-----Ala	297	::: :::
Db	1069 TCACCACTGCTTCAGCTTCAGCTTCAGCTCAACAGTGCCTCACGACAGATTCACAGC	1128	::: :::
Oy	298 GlyGlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAla	317	::: :::
Db	1129 TCTGATTCGATCAACAGAGGCGTTCGATGATGATCAACAGATGATCGATCGATCAACA	1188	::: :::
Oy	318 ValValGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAla	337	::: :::
Db	1189 AGCAACAAGTCTTCAGGCTCAGCAAGATATTCAGGCTGATTCGCAATCGGCATCAACAGTGGC	1248	::: :::
Oy	338 IleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle	357	::: :::







```

Db 4688 GCTTCAGCAAGTACC-----TCAGCGCTCGAATCA 4717
Oy 42 LysGlnIleGlnInThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIle 61
Db 4718 GCATCAACAACTGGCGGCTTCAGCATCAACGAGTCATCAGCTTCAGCATCAACAACT 4777
Oy 62 AlaGlyAspSerGlyLysAspLysThrSerThrThrLysThrGluThrAlaProGln 81
Db 4778 GCTTCAGCTTCAGCAAGTACAGTCGCGCTTCAGCATCAACGAGTCGCTTCAGCTCA 4837
Oy 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGly 101
Db 4838 GCCTCAACAGGCTCGCTCGAATCCGATCAACAGTCGCTCGCTTCAGCAAGCACAGT 4897
Oy 102 ValSerGlyAlaAlaAlaThr-----ThrAlaSerAsnThrAlaThrIle 117
Db 4898 GCTTCGCTTCAGCATCAACGAGTCGCTTCAGTCATCAACGAGTCGCTTCAGCTCA 4957
Oy 118 AlaMetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThrLysGluSerLeu 137
Db 4958 GCAACGACATCAGCTTCGAAATCTGCATCAACCACTGGCTGCTCCGACATCAACAAAGC 5017
Oy 138 GlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSer 157
Db 5018 GCCTCGGCTCAGCAAGTACAAAGTCTTCAGCTCAGCATCAACCAAGTCATCAGCTTCA 5077
Oy 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProLysValThr 177
Db 5078 GCCTCAACAAAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 5137
Oy 178 ProArgSerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGly 197
Db 5138 GCGTCAGCTTCAGCAACGACAAAGTCGTCAGCTTCAGCATCAACGAGTCGCTTCG 5197
Oy 198 GlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThr 217
Db 5198 GCATCAACAAAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCT 5251
Oy 218 AsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgLysLysGln 237
Db 5252 ACT-----AGTCATCAGCATCAGCATCAACCAAGTCATCAGCTTCAGCA 5296
Oy 238 GlnMetLysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsn 257
Db 5297 AGTATCTCAGCTTCGAAATCGGCATCAACGAGTCATCAGCTTCAGCATCAACGAGTGA 5356
Oy 258 ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePhe 277
Db 5357 TCGGCTTCAGCTCAACGAGTCATCAGCTTCAGCAACGACCAAGTCGCTTCAGCA 5416
Oy 278 ThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAla 297
Db 5417 TCACGAGTGCCTCAGCTTCAGCAAGTATCTCAGCTTCGAAATCGGCATCAACGAGTGG 5476
Oy 298 GlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 317
Db 5477 TCAGGCTTCAGCAAGTACGATCGCTTCAGCAACGAGTCGCTTCAGCA 5536
Oy 318 ValValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla 337
Db 5537 TCACACAGTGCCTCAGCTTCAGCAAGTATCTCAGCTTCGAAATCGGCATCAACGAGTGG 5596
Oy 338 IleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
Db 5597 ---TCAGCTTCAGCAAGTACGATCGCTTCAGCAACGAGTCGCTTCAGCA 5653
Oy 358 AlaLysAlaIleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAla 377
Db 5654 GCAATACAGGCGCTTCAGCTTCAGCAACGAGTCGCTTCAGCA 5698
Oy 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrVal 397
Db 5699 ---TCACCAAGTACCAAGCGCTTCAGCTTCAGCAACGAGTCGCTTCAGCA 5740

```

```

Oy 398 GlyValGlyValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu 417
Db 5741 TCAGCTTCAGCAAGTACGATCGCTTCAGCTTCAGCATCAACGAGTCGCTTCAGCTTCG 5797
Oy 418 SerGlnMetGlnInAsnValAlaGlnPheGln---LysGlnValGlyLysLeuGlnAla 436
Db 5798 ---CCTCAACAAAGTGCATCGCTTCAGCTTCAGCAACGAGTCGCTTCAGCA 5848
Oy 437 Ala-----AlaAspMetIleSerMetPheThrGlnPheTrpGln 449
Db 5849 GCACCAAGTGCCTCGGCTTCAGCAACGAGTCGCTTCAGCTTCAGCTTCAGCTTCAGCA 5902
Oy 450 GlnAlaSerLysIleAlaSer-----LysGln 458
Db 5903 GT-GGCTCAGCTTCAGCAAGTACATCAGTTCAATTCAGCAACCAATTCGAACTCAGCA 5961
Oy 459 ThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLys----- 471
Db 5962 GTTGGAAATACTTCTGATCGACAGCTAAATCCAAAAAGAAATTCCTTAATACAGTACT 6021
Oy 472 -----LysGlyAlaGlnIleLeuLysAlaTrpAlaAlaIleSerGly 485
Db 6022 GAGTCGTCATGATGATGTGTACTTGAGTTCAGCAAGCTTCAGCAAGT 6072

RESULT 13
AAV52496
ID AAV52496 standard; DNA; 4483 BP.
XX
AC AAV52496;
XX
DT 23-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae genome fragment SEQ ID NO:363.
XX
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KM computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS Streptococcus pneumoniae.
XX
PN W09818931-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US19588.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX Kunsch CA, Rosen CA.
XX
DR WPI: 1998-272225/24.
XX
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
PS
PS Claim 1: Page 1376-1378; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
XX the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52574)
XX recorded on it, or a representative fragment or a sequence at least 95%
XX identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
XX SEQ ID NO:1 to 391 (AAV52134 to AAV52574) are genomic fragments from
XX Streptococcus pneumoniae. The present invention also describes an
XX isolated nucleic acid molecule encoding a homologue of any of the
XX fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the
XX nucleic acid molecule is produced by a process comprising: (a) screening
XX a genomic DNA library using as a probe a target sequence defined by any
XX of the sequences in SEQ ID NO:1 to 391, identifying members of the

```



```
OY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
DB 484 ATGCTATATTCATCTTCTTCAGAGCCTGACATCAAAAAATATCATGCTCAAGTCTG 543
OY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
DB 544 ACATGACACCCCGAGGGGCGGCCCAACAAAGATGCTGCGGACGAAACGAAAGCA 603
OY 44 IleGlnIleThrArgGlnGlyLysAsnThrGlnMetGlnMetGlnSerAspAlaThrIleAlaGly 63
DB 604 ATACAGCAACACGCTCAGGCTTAACACACTGATGATGAAAGGATGCCACTTGTGCTG 663
OY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
DB 664 GCTTCGGAAGAACAAACTCTCCACTACAAAACGAAACAGCTCCACACAGGGA 723
OY 84 ValAlaIleAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrLysValSer 103
DB 724 GTTGGCTGGGAAAGATCTCCAGAAAGTCAAAAGCGGCTGATACCTGAGATCA 783
OY 104 GlValAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 784 GGAGCGGCTGCTACTACAGCATCAAACTACTGCAACAAAATTCCTATGACAGCTTATT 843
OY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
DB 844 GAAGAGCGCAGCAAAAGATGATGAGTCTACCTTAGAGTCACTCAAAAGCCTCAGTGGCG 903
OY 144 GlnMetLysGlnValGlnAlaValValAlaIleAlaSerGlyLysSerSerGlySer 163
DB 904 CAAATGAAAGAGTCAAGACGGTGTGTGTGCTGCCCTCAGGAAAGGTTGCGGTTCC 963
OY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183
DB 964 GCAAAATTTGGAACACCTAGCTCCCAAGCCCGGGGTACACCAAGATCAGAGTTATC 1023
OY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
DB 1024 GAAATCGAGCTCGCGCTCTTAAGCAATTCAGACATGTGGAGAAACCAAAATCTGCC 1083
OY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeuGln 223
DB 1084 TTATCTTAACCTATGCAAGTACACAGCAAGCAGACCAAACTAATAGCTTAGAA 1143
OY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnMetLysAlaIleGln 243
DB 1144 AAGCAAGCATTAATAATTCATTAAGAACAGAAAGATACCAGAGATGAAGGCTGCCGA 1203
OY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
DB 1204 CAGAAGCTTAAGATCTCGAAGGAAATGATGATGATGATGATGATGATGATGATGATG 1260
RESULT 11
AAZ01425/c
ID AAZ01425 standard; DNA; 1038602 BP.
AC AAZ01425;
XX
XX
XX 07-OCT-1999 (first entry)
XX
XX Complete genome sequence of Chlamydia trachomatis.
XX
XX Vaccine, eye disease; conventional trachoma; nonendemic trachoma;
XX paratrachoma; inclusion conjunctivitis; genital disease; perithelphalitis;
XX nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
XX Bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX Chlamydia trachomatis.
XX
XX OS
XX PN MO9928475-A2.
XX PD 10-JUN-1999.
XX
```

```
PF 27-NOV-1998; 98W0-IB01939.
XX
XX 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
XX (GBST ) GENSET.
XX
XX PI Griffiths R;
XX
XX WPI: 1999-371125/31.
XX
XX Genome sequence of Chlamydia trachomatis
XX
XX Claim 1: Page 373-656; 1755pp; English.
XX
XX The present sequence represents the complete genome of Chlamydia
XX trachomatis. Open reading frames (ORFs) of the genome encode
XX polypeptides AA36754-37949. The polypeptides can be used as vaccines
XX against Chlamydia trachomatis. Antisense and ribozyme sequences can also
XX be used to control growth of the microorganism. Chlamydia trachomatis is
XX responsible for a large number of diseases, e.g. eye diseases such as
XX conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
XX conjunctivitis; genital diseases such as nongonococcal urethritis,
XX epididymitis, cervicitis, salpingitis, perithelphalitis, Bartholinitis;
XX pneumopathy in breast feeding infants; and venereal
XX lymphogranulomatosis. The polypeptides of the invention may be of use in
XX treating these diseases.
XX
XX Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other:
SQ
Alignment Scores:
Pred. No.: 1,286-76 Length: 1038602
Score: 1195.50 Matches: 266
Percent Similarity: 69.80% Conservative: 83
Best Local Similarity: 53.20% Mismatches: 124
Query Match: 50,468 Indels: 28
DB: Gaps: 11
US-09-889-314-2 (1-496) x AAZ01425 (1-1038602)
OY 4 MetSerIleSerSerSerSerGlyProAsp-----AsnGlnLysAsnIleMetSerGln 21
DB 863525 ATGTCCTTTCATCTTCTTCAGAGCCTGACATGACCACTTAAGATCTCTGTCGCA 863466
OY 22 ValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThr 41
DB 863465 GTCATAGCTTCGACTCTTAAGCGCTTCTTAATGCAATTAATTAACGACATCAGTT 863406
OY 42 LysGlnIleGlnIleThrArgGlnGlnLys---AsnThrGlnMetGlnSerAspAlaThr 60
DB 863405 AAGCAAGTTCACAGACGAGACGAAATAATCGCATGACCTTAAGCATGAGAGCGATGCGCT 863346
OY 61 IleAlaGlyAlaSerLysAspLysThrSerSerThrThrLys-----ThrGln 77
DB 863345 GTTGGCGGAACGCTGTAAGAAAGATGCGGCACTTCTTCAATTAAGACAGCAAGA 863286
OY 78 ThrAlaProGlnGlnGlyValAlaIleAlaGlyLysGlnSerSerGlnSerGlnLysAlaGly 97
DB 863285 CTTATAGACGACGAAAGATTAGCTGACGGAAGAAGAACT-----GCATCT 863241
OY 98 AlaAspThrGlyValSerGlyAlaAlaIleAlaThrThrAlaSerAsnThrAlaThrLys--- 116
DB 863240 GCGGAT-----GGACATTCCTTAACCCAAAGCGCATCTTAAGCA 863202
OY 117 IleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGln---SerThrLeuGln 135
DB 863201 GCTAGCTCGCAACATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 863142
OY 136 SerLeuGlnSerLeuSerAlaIleAlaGlnMetLysGlnValGlnAlaValAlaAlaAla 155
DB 863141 TCGTTGTGATCTGTAGATGCGACGATCTTAAGAAATTCAAAGATCGATGCTGCGCT 863082
```



CC with the expression vector pAD431 to give pCPN533alpha. The  
CC plasmid was used to transform an E. coli host, which was cultured  
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A  
CC and primers and probes derived from its DNA can be used in assays  
CC for the detection of polypeptide A antibodies and DNA,  
CC respectively, useful in the diagnosis of C. pneumoniae infection.

XX Sequence 813 BP; 286 A; 181 C; 187 G; 159 T; 0 other:

## Alignment Scores:

Pred. No.:	5,58e-84	Length:	813
Score:	1245.00	Matches:	259
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.55%	Indels:	0
DB:	17	Gaps:	0

US-09-889-314-2 (1-496) x AAT14613 (1-813)

```
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
Db 1 ATGCTATTTCATCTTCTTCAGACCTGACATCAAAAAAATATCATGTCTCAAGTCTGTG 60
QY 24 ThrSerThrProGlnGlnValProGlnLysLeuSerGlyAsnGlnThrLysGln 43
Db 61 ACATCGACCCCGAGGGGCTCCCAACACATGAGTCTGTGGCAACGAAACCAACGCA 120
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGly 63
Db 121 ATACAGCAACACGTCAGGTAAACACTGAGATGGAACGATGCGCATATTGTCTGT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
Db 181 GCTTCGAAAAAGACAAACTCTCGACATCAAAACAGAACAGCGTCCACACAGGGA 240
QY 84 ValAlaIleGlyLysGluSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 241 GTTGCTGCTGGGAAGAAATCTCTCAAAAGTCAAAAGCAGGTGCGATCGAGTATCA 300
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 301 GGAGCGGCTGTACTACACGATCAAAATCTGCAACAAAATTTGCTATCTACAGACTCT 360
QY 124 GlnGluAlaSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
Db 361 GAGAGGCGAGCAAAAGATGAGTCTTACCTTAGAGTCACTTCAAGCCTCAGTCCGCG 420
QY 144 GlnMetLysGluValAlaGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 421 CAAATGAAAGAGTCGAGGCGGTGTGTGCTGCCCTCTCAGGGAAAGTTCCGGTCC 480
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
Db 481 GCAAAATTTGGAAACCTGAGTCCCAAGCCGGGGTGACCCACAGATCAAGGTTTATC 540
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyAlaAlaThrLysSerAla 203
Db 541 GAAATCGGACTCGCGCTTCTTAAAGCAATTCAGACATTGGGAGAACCCAAAAATCTGCC 600
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeu 223
Db 601 TTATCTAACTATGCAAGTACACACACACACACCAAAACCAAAATTAAGTCTAGAA 660
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleMetLysAlaAlaGlu 243
Db 661 AAGCAAGCATATAAAATCATTAAGAAAGAGAAATACCAAGAGTGAAGCGCTGCCGA 720
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
Db 721 CAGAAGTCTTAAGATCTCGAAGAACAAATGATACGTGCTATCTGATGATCGCG 777
```

RESULT 10  
AAT14619

ID AAT14619 standard; DNA; 1296 BP.

XX AAT14619;

XX 11-NOV-1996 (first entry)

XX DHFR/C. pneumoniae antigen variant fusion protein DNA.

XX Polypeptide antigen; strain YK41; plasmid; probe;

XX pCPN533T; primer; assay; detection; antibody; diagnosis;

XX infection; fusion protein; dihydrofolate reductase; DHFR;

XX variant; Chlamydia pneumoniae; ds.

XX Synthetic.

XX Key Location/Qualifiers

XX CDS 1..1296

XX /tag= a

XX /note= "STOP codon absent"

XX W09609320-A1.

XX 28-MAR-1996.

XX 20-SEP-1995; 95MO-JP01896.

XX 28-APR-1995; 95JP-0106011.

XX 20-SEP-1994; 94JP-0224711.

XX 28-APR-1995; 95JP-0106006.

XX 28-APR-1995; 95JP-0106008.

XX 28-APR-1995; 95JP-0106009.

XX 28-APR-1995; 95JP-0106010.

XX (HITB ) HITACHI CHEM CO LTD.

XX Iutsu H, Matsumoto A, Obara K;

XX WPI; 1996-188399/19.

XX P-PSDB; AAR94585.

XX Recombinant Chlamydia pneumoniae antigen and antibodies to it

XX used for detection and assay of C. pneumoniae e.g. in clinical

XX diagnosis

XX Claim 26; Pages 103-107; 128pp; Japanese.

XX The present sequence encodes the dihydrofolate reductase (DHFR)/

XX C. pneumoniae polypeptide antigen variant fusion protein.

XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted

XX to prep. a lambda gt11 DNA library. The library was then screened

XX with an anti-YK41 monoclonal antibody (Mab), which was prepd. by

XX fusing spleen cells from a mouse infected with YK41 with myeloma

XX p3/NSI/1-Ag4-1 to produce a Mab expressing hybridoma. The DNA

XX obtd. was then fused with DHFR DNA and the expression vector

XX pAD431 to give pCPN533T. The plasmid was used to transform an

XX E. coli host, which was cultured to give an antigenic polypeptide

XX fusion protein. The fusion protein and primers and probes derived

XX from its DNA can be used in assays for the detection of the

XX antigenic polypeptide antibodies and DNA, respectively, useful in

XX the diagnosis of C. pneumoniae infection.

XX Sequence 1296 BP; 407 A; 300 C; 324 G; 265 T; 0 other:

XX Alignment Scores:

Pred. No.:	9,68e-84	Length:	1296
Score:	1245.00	Matches:	259
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.55%	Indels:	0
DB:	17	Gaps:	0

US-09-889-314-2 (1-496) x AAT14619 (1-1296)



```

XX  JP08304404-A.
XX
XX  22-NOV-1996.
XX
XX  28-APR-1995; 95JP-0106014.
XX
XX  28-APR-1995; 95JP-0106014.
XX
XX  (HITB ) HITACHI CHEM CO LTD.
XX
XX  WPI; 1997-056178/06.
XX  P-PSDB; AAM01743.
XX
XX  Detection and determination of anti-Chlamydia pneumoniae antibody -
XX  using the polypeptide C as the antigen
XX
XX  Example 3; Page 17-18; 18pp; Japanese.
XX
XX  The gene encodes the Chlamydia pneumoniae 53 kDa antigen (AAM01743). A
XX  method for the detection and determination of anti-C. pneumoniae
XX  antibodies in a sample comprises using at least 5 consecutive amino
XX  acids of the polypeptide C 73 kDa antigen (AAM01742).
XX
XX  Sequence 1048 BP; 360 A; 241 C; 231 G; 216 T; 0 other;
SQ

```

Alignment Scores:

Pred. No.:	4,09e-85	Length:	1048
Score:	1262.00	Matches:	262
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	53.27%	Indels:	0
DB:	18	Gaps:	0

US-09-889-314-2 (1-496) x AAT59311 (1-1048)

```

OY  1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
    |||||
DB  227 GATACAAACATGCTTTCATCTTCTTCAGACGCTGACACATATAAATATCATGTCT 286
OY  21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40
    |||||
DB  287 CAAGTCTGACATCGACACCCCGCGCGTCCCAACAAGATAGCTGCTGCGACAGAA 346
OY  41 ThrGlyGlnIleGlnIleThrArgGlnGlyLysAsnThrGlnMetLysSerAspAlaThr 60
    |||||
DB  347 ACCGAACCAATATACAGAAACGCTCAGGTAATAAACACTGAGATGGAAGCCGATGCCACT 406
OY  61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
    |||||
DB  407 ATTGCTGCTGCTTCTGGAAGCAAACTCTCTCGACTACCAAAAACAGAACAGCTCCA 466
OY  81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100
    |||||
DB  467 CAACAGGAGATTGCTGCTGGAAGAAATCTCTCAGAAAGTCAAAAAGCAGAGTCTGATAC 526
OY  101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
    |||||
DB  527 GGAGTATCAGAGAGCGGCTCTACTACAGCATCAAAATATGCAACAAATGCTATGCA 586
OY  121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerLeu 140
    |||||
DB  587 ACCCTATTGTAAGAGCGGCAACAAATATGAGTCTTACCTTAGAGTCACTTCAAAAGCTC 646
OY  141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSerGlyLysSer 160
    |||||
DB  647 AGTGGCGCCCAATGAAAGAGCGAGCGGTGTGTGCTGCCCTCAGAGGAAAG 706
OY  161 SerGlySerAlaLysLeuGlnThrProGlnLysProGlyValThrProAlaGlySer 180
    |||||
DB  707 TCGGGTTCCGCAAAATTTGGAAACCTCGAGCTCCCAAGGCCGGGTGAGACCAACATCA 766
OY  181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnIleThrLeuGlnGlyAlaThr 200
    |||||

```

```

DB  767 GAGGTATCGAATCGAGCTCGCGCTTGCTTAAGCAATTCAGATGTGGAGAACGCACA 826
    |||||
OY  201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrLysLeu 220
    |||||
DB  827 AATCTGCTTATCTAATCATGCAAGTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 886
OY  221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGluThrGlnGlnMetLys 240
    |||||
DB  887 GGTCTAGAAAGCAAGCGATTAATAATCGATTAAGAGAGAGAGATACCAAGAGATGAAG 946
OY  241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
    |||||
DB  947 GCTGCCGCAACAGACTCTAAGATCTCGAAGGACAAATGATCTGATCTGATGATG 1006
OY  261 IleAla 262
    |||||
DB  1007 ATCGCG 1012

```

RESULT 9

AAT14613

ID AAT14613 standard; DNA; 813 BP.

AC AAT14613;

DT 07-NOV-1996 (first entry)

DE C. pneumoniae polypeptide antigen (polypeptide A) DNA variant.

XX Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;

KW PCP533alpha; primer; assay; detection; antibody; diagnosis;

KW Infection; variant; ds.

XX Chlamydia pneumoniae.

OS

XX

XX Key Location/Qualifiers

FT CDS 1..812

FT /\*tag= a

FT /note= "STOP codon absent"

XX

XX W09609320-A1.

XX

XX 28-MAR-1996.

XX

XX 20-SEP-1995; 95WO-JP01896.

XX

XX 28-APR-1995; 95JP-0106011.

XX

XX 20-SEP-1994; 94JP-0224711.

XX

XX 28-APR-1995; 95JP-0106006.

XX

XX 28-APR-1995; 95JP-0106008.

XX

XX 28-APR-1995; 95JP-0106009.

XX

XX 28-APR-1995; 95JP-0106010.

XX

XX (HITB ) HITACHI CHEM CO LTD.

PI

XX Izutsu H, Matsumoto A, Obara K;

XX

XX WPI; 1996-188399/19.

XX

XX P-PSDB; AAR94580.

XX

XX Recombinant Chlamydia pneumoniae antigen and antibodies to it

PT used for detection and assay of C. pneumoniae e.g. in clinical

PT diagnosis

XX

XX Claim 10; Pages 64-66; 128pp; Japanese.

XX

XX The present sequence encodes a variant of the C. pneumoniae

XX polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was

XX cultured and genomic DNA extracted to prep. a lambda gt11 DNA

XX library. The library was then screened with an anti-YK41

XX monoclonal antibody (Mab), which was prep. by fusing spleen cells

XX from a mouse infected with YK41 with myeloma P3/NSI/1-Ag4-1 to

XX produce a Mab expressing hydridoma. The DNA obtd. was then fused

AA14622  
ID AA14622 standard; DNA; 1048 BP.  
AC AA14622;  
XX  
DT 11-NOV-1996 (first entry)  
DE C. pneumoniae polypeptide antigen (polypeptide A) DNA clone 53-35.  
XX  
KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;  
KW pCPM533alpha; primer; assay; detection; antibody; diagnosis;  
KM Infection; clone; ds.  
XX  
OS Chlamydia pneumoniae.  
XX  
FH Key Location/Qualifiers  
FT CDS 236..1012  
FT /tag= a  
FT /note= "STOP codon absent"  
XX  
PN WC0609320-A1.  
XX  
PD 28-MAR-1996.  
XX  
PE 20-SEP-1995; 95WO-JP01896.  
XX  
PR 28-APR-1995; 95JP-0106011.  
PR 20-SEP-1994; 94JP-0224711.  
PR 28-APR-1995; 95JP-0106006.  
PR 28-APR-1995; 95JP-0106008.  
PR 28-APR-1995; 95JP-0106009.  
PR 28-APR-1995; 95JP-0106010.  
XX  
PA (HITB) HITACHI CHEM CO LTD.  
XX  
PI Izutsu H, Matsumoto A, Obara K;  
XX  
DR WPI; 1996-188399/19.  
DR P-PSDB; AAR94586.  
XX  
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it  
PT used for detection and assay of C. pneumoniae e.g. in clinical  
PT diagnosis  
XX  
PS Example 1; Pages 75-77; 128pp; Japanese.  
XX  
XX The present sequence encodes the C. pneumoniae polypeptide antigen  
CC polypeptide A clone, 53-35. C. pneumoniae strain YK41 was  
CC cultured and genomic DNA extracted to prep. a lambda gill DNA  
CC library. The library was then screened with an anti-YK41  
CC monoclonal antibody (Mab), which was prepd. by fusing spleen cells  
CC from a mouse infected with YK41 with myeloma P3/NS1/I-Ag4-1 to  
CC produce a Mab expressing hybridoma. The DNA obcd. was then fused  
CC with the expression vector pAD431 to give pCPM533alpha. The  
CC plasmid was used to transform an E. coli host, which was cultured  
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A  
CC and primers and probes derived from its DNA can be used in assays  
CC for the detection of polypeptide A antibodies and DNA.  
CC respectively, useful in the diagnosis of C. pneumoniae infection.  
XX  
SQ Sequence 1048 BP; 360 A; 241 C; 231 G; 216 T; 0 other;

Alignment Scores:  
Pred. No.: 4 09e-85 Length: 1048  
Score: 1262.00 Matches: 262  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.27% Indels: 0  
DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AA14622 (1-1048)

OY 1 AsptHrAmMetSerIleSerSerSergIlyProAspAsnGlnIlyAsnIleMetSer 20

DB 227 GATACAAACATGCTATTTCATCTTCTTCAGACCTCGAACATCAAAAAATATCATGTCT 286  
OY 21 GlnValIleuThrSerThrProGlnIlyValProGlnIlyAspIlySerGlyAsnGlu 40  
DB 287 CAAGTCTCAGATCGACACCCAGGGCGTGGCCCAAGATTAAGCTGTCTGGCAACGAA 346  
OY 41 ThrLysGlnIleGlnIlnThrArgGlnIlyAsnThrGluMetGluSerAspAlaThr 60  
DB 347 ACCAAGCAATATACGCAAAACGCTGAGGTAAACACAGTGAATGGAAGCCATGCCACT 406  
OY 61 IleAlaGlyAlaSerGlyIlyAspIlyThrSerSerThrThrLysThrGluThrAlaPro 80  
DB 407 ATTGCTGGTCTCTTGAAAAAGCAAACTCTCTCATCAAAAAACAGAAACAGCTCCA 466  
OY 81 GlnGlnIlyValAlaAlaGlyIlyGluSerSerGluSerGlnIlyAlaIlyAspThr 100  
DB 467 CAACAGGAGTGTCTGCTGGGAAAGATCTCTGAGAAAGTCAAAAGGACAGTCTGATTA 526  
OY 101 GlyValSerGlyAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
DB 527 GGAGTATCAGAGCGGCTGCTACTACAGCATCAATATCTGCACAAATAATGCTATCGAG 586  
OY 121 ThrSerIleGlnIlyAlaSerIlySerMetGluSerThrLeuGluSerLeuGlnSerLeu 140  
DB 587 ACCCTCTTGAAGAGGAGGACCAAAAGATAGTACGTCTTACGTACACTTCAAAAGCTTC 646  
OY 141 SerAlaAlaGlnMetIlyGluValGluAlaValAlaValAlaAlaIleuSerGlyLysSer 160  
DB 647 AGTCCCGCCGCAAAATGAAAGAGTGAAGCGGTGTGTGCTCCCTCTCAGGGAAAAAGT 706  
OY 161 SerGlySerAlaIlyLeuGluIlyThrProGluLeuProIlyProGlyValThrProArgSer 180  
DB 707 TCGGGTTCGGCAAAATTCGAAACACCTGACCTCCCAAGCCGGGGTGAACCAAGATCA 766  
OY 181 GlnValIleGlnIleGlnIlyLeuAlaIleuAlaIlyAlaIleGlnThrLeuGlnIlyAlaThr 200  
DB 767 GAGGTATCGAAATCGGACATCGCCCTTGAAGCAATTCAGCATTTGGGAGAGGCACA 826  
OY 201 LysSerAlaIleuSerAsnIlyAlaSerThrGlnAlaGlnAlaAspGlnIlyAsnIlyLeu 220  
DB 827 AAATCTGCTTATCTTAATCTATGCAATGACACACACACACACACACACACAAATTAAC 886  
OY 221 GlyLeuGluIlyGlnAlaIleIlyIleAspIlyGluArgGluIlyGlnIlyMetLys 240  
DB 887 GGCTCAGAAAGCAAGCGATTAATGATTAACACAGAGAGATATCCAAAGAGATGAAG 946  
OY 241 AlaAlaGlnIlyIlySerIlyAspLeuGlnIlyThrMetAspThrValAsnThrValMet 260  
DB 947 GCTGCCGACAGAAAGTCTAAAGATCTGAAAGAACATGATGATCTGATCTGATGATG 1006  
OY 261 IleAla 262  
DB 1007 ATCGCG 1012  
RESULT 8  
ID AA159311 standard; DNA; 1048 BP.  
AC AA159311;  
XX  
DT 22-APR-1997 (first entry)  
DE  
XX  
XX DNA encoding C. pneumoniae 53 kDa antigen.  
XX  
KW antigen; antibody; detection; determination; epitope; ds.  
XX  
OS Chlamydia pneumoniae.  
XX  
FH Key Location/Qualifiers  
FT CDS 236..1012  
FT /tag= a  
FT /note= "no stop codon present"



```

Db      811 CAACTTCTGACATGACACCCCGGCGTCCCAACAAATAGCTGTGCAACGAA 870
Qy      41 ThrlysglnileglnlnthraraglnlyllysasnThrslumetgluseraspalaThr 60
      871 ACGAAGCAAAATACGCAACACGTCAGGTAATAACACTGAGATGGAACGATGCCACT 930
Qy      61 llaalaglyalaserglyllyaspllysthrSerSerThrThrlysthrlythrlnalapro 80
      931 ATTGCTGGTCTCTGGAAGAACAAACTCTCGACTACAAAAACAGAAACAGCTCCA 990
Qy      81 GlnGlnGlyAlaAlaAlaGlyllysgluserSerGlnuserGlnlysalaglyalaaspthr 100
      991 CAACAGGAGTTGCTGCTGGGAAAGAAATCTCAGAAAGTCAAAAGCGAGGCTGATACT 1050
Qy      101 GlyAlaSerGlyAlaAlaAlaAlaThrThrAlaSerasnThrAlaThrlylAlaMetGln 120
      1051 GGAGATACAGAGCGCGTCTACTACACATCAATACTGCACAAATAATTGCTATGCGAG 1110
Qy      121 ThrSerlileglnGlnAlaSerlySerMetGlnSerThrleugluserleuglnSerleu 140
      1111 ACCCTATATGAGAGGCGGCAAGCAATGAGTGTACTTACCTTACAGTCAACTCAAGCCTC 1170
Qy      141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaSerGlyLysSer 160
      1171 ACTGCCGCGCAAAATGAAGAGTCGAAAGCGGTTGTTGCTGCCCTCTCAGGGAAGAGT 1230
Qy      161 SerGlySerAlaLysleuglnurthrProgluleuProLysProGlyValThrProargSer 180
      1231 TCGGGTCCGCAAAATGGAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290
Qy      181 GlnValileglnileglnleuAlaAlaAlaLysAlaileglnThrleuglyGlnAlaThr 200
      1291 GAGGTATGAAATGAGATCGCGCTCGCTCTTAAGCAATTCAGACATTCGAGAACACCA 1350
Qy      201 LysSerAlaLeuSerasnThrAlaSerThrGlnAlaAlaSerGlnThrAsnLysleu 220
      1351 AATTCGCTTATCTATCTATGCAATGCAATGCAACAGCAAGCAAGCAACCAATTAACCA 1410
Qy      221 GlyleuglnLysGlnAlaAlaLysIleAspLysGlnArgGlnLysGlnleuglnleuLys 240
      1411 GGTCTGAAAGCAAGCAAGCAATTAATGCAATGAAGCAAGCAAGCAATGAAGATGAAG 1470
Qy      241 AlaAlaGlnLysSerLysAspLeuglnGlyThrMetAspThrValAsnThrValMet 260
      1471 GGTGCCGAAACGAAATCTAAAGATCTCGAAGCAATGATCTGTCACTACTGTGATG 1530
Qy      261 llaAlaValSerValAlaAlaIleThrVal-----lIeSer 271
      1531 ATCCGCAAA-GGGGTTGCAATTTGCCATGCGGCGCTTAATTAACTGAGAGATCCAG 1589
Qy      272 lIeValAlaAlaIlePheThrCysGlyAlaGlyleuAlaAlaGlyleu-----AlaAlaGly 289
      1590 ATCTAATCTGATGATCTCTCTACGCGGACGATGTCGCGCATCCCGGCGCCACAGGT 1649
Qy      290 AlaAlaValAlaGlyAlaAla----- 295
      1650 GCGGTTCCTGCGGCTTA-TATCCGCGACATCCAGATGGGAGATCGGCGTCCGCACTT 1708
Qy      296 -----AlaAlaGlyGlyAla 300
      1709 CGGGCTCATGAGCGCTTGTTCGGCGTGGGATGATGTCGAGCGCCGCGGGGAGCTG 1768
Qy      301 AlaAlaGlyAla-----AlaAlaAla 306
      1769 TTGGGCGCCATCTCTTCATGACACCATTCCTTCGCGGCGGTGCTCAACGCGCTCAAC 1828
Qy      306 ----- 306
      1829 CTACTACTGGGCTGCTCTCTAATGACAGAGTCGATAGGAGAGCGTCGACCGATGCC 1888
Qy      307 -----ThThValAla 310
      111 111111

```

```

Db      1889 TTGAGAGCCTTCAACCCAGTCACTCTTCGCGTGGGCGGCGGCAATGACTATGCTGCC 1948
Qy      311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValAlaLysGlnAlaValIleThrAlaVal 330
      1949 GCACATTATGACTGCTCTCTTATCATGCACTGTAAGACAG-----LysAla 340
Qy      331 ArgGlnAlaIleThrAlaAlaIle----- 340
      1991 -----GTCCCGGCAAGCGCTTGGGTCATTTCGCGGAGAACCGCTTTCGCTGAGC 2041
Qy      341 AlaValLysSerGlylIleLysAlaPheIleThrThrleuValLysAlaIleAlaLysAla 360
      2042 GCGACGATGATGCGCTGCTGCGCTTGGGATTCGGAATCTTGACAGCCCTGCTCAAGCC 2101
Qy      361 lIeSerLysGlylIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
      2102 TTGCTACGTGCTCCGCGC-----ACCAAGGTTTC-2131
Qy      381 ProLysLeuSerLysValIleSerSerleu----- 390
      2132 GCGGAGAACGAGCGCAATTCGCGCGCATGCGGCGGCGGCGCTGAGCTTGTGCTG 2191
Qy      391 -----ThrSerlystrpValThrValGlyValGlyValAlaAlaAlaProAla 407
      2192 GCGTTCGCGAGCGGAGGCTGAGTGCCTTCCCATTAATGATTCCTGCTTCGCGCGCG 2251
Qy      408 LeuGlyLysGlylIleMetGlnMetGlnLeuSerGlnMetGlnMetGlnAsnValAlaGlnPhe 427
      2252 ATCGGGATGCGCGCGCTTGCAGGCGCATGCTGTCCAGGACAGTAGATACGACCATCGAG 2311
Qy      428 GlnLysGlnValGlyLysLeuGlnAlaAlaAlaAlaAspMetIleSerMet 443
      2312 CAGCTTCAA---GGATGCTGCGGCGCTTACCAAGCTACTTGCATC 2356
Db      2312 CAGCTTCAA---GGATGCTGCGGCGCTTACCAAGCTACTTGCATC 2356

RESULT 6
ID      AAT33893
AC      AAT33893;
CC      11-NOV-1996 (first entry)
DE      C. pneumoniae antigen (polypeptide A) expression vector.
KW      Polypeptide antigen; polypeptide A; strain Y41; plasmid; probe;
      PCPN533a1pha; primer; assay; detection; antibody; diagnosis;
      infection; Chlamydia pneumoniae; ds.
OS      Synthetic.
FH      Key
FT      CDS
      Location/Qualifiers
      481..2428
      /*tag= a
WO9609320-A1.
28-MAR-1996.
20-SEP-1995; 95WO-JP01896.
28-APR-1995; 95JP-0106011.
20-SEP-1994; 94JP-0224711.
28-APR-1995; 95JP-0106006.
28-APR-1995; 95JP-0106008.
28-APR-1995; 95JP-0106009.
28-APR-1995; 95JP-0106010.
(HITB ) HITACHI CHEM CO LTD.
Izutsu H, Matsumoto A, Obara K;
WPI; 1996-188399/19.
P-PDB; AAR94579.

```

Qy	164	AllylseneuLuthrProgiueuPolysProgiyalThProarSeriLualile	183
Db	481	GC AAAATGGAAACACCTGAGCTCCCAAGGCCGGGGTGCACCAAGATCAGAGTTATC	540
Qy	184	GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla	203
Db	541	GAATCGACTCGCGCTTGCTTAAGCAATTCAAGACTTGGGAGAACCAAAAATCTGCC	600
Qy	204	LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu	223
Db	601	TTATCTAACTATCCAGATACACAAAGCACAAGCAGACCAAAACAAATAACTAGCTGTGAA	660
Qy	224	LysGlnAlaIleLysIleAspLysGluATGGGJugIuTyrGlnGluMetLysAlaGlu	243
Db	661	AAGCAACGATAAATAATCGATTAAAGACAGAAAGATATCCAAAGATGAAAGGCTGGCAA	720
Qy	244	GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaVal	263
Db	721	CAGAAAGCTAAAGATCTCGAAGGACAAATGATGACTGTCATATCTGATGATCGCGGTT	780
Qy	264	SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu	283
Db	781	TCGTGTCCATTCACGTTATTTCTATTGTGGCTATTTTTACATGGCGAGCTGACTC	840
Qy	284	AlaGlyLeuAlaIleGlyAlaAlaValGlyAlaAlaIleGlyAlaGlyAlaGlyAla	303
Db	841	GCTGACTCGCTCGCGAGCTGCTGTACGTGCAAGCGGCAAGCTGGAGGTGCAGCGACT	900
Qy	304	AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys	323
Db	901	GCTGCCCAACCAACGGTAGCACCAAAATTACAGTTCAAGCTGTTGTCCAAAGCGGTGAA	960
Qy	324	GlnAlaValIlePheAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys	343
Db	961	CMACTGTATTCACAGCTGTGCACAGCAAGGATCACCGCGCTATTAAGCGGCTGTCAA	1020
Qy	344	SerGlyIleLysAlaPheIleLysThrIleuValLysAlaIleAlaLysAlaIleSerLys	363
Db	1021	TCGTGATTAATAAGCATTTATCAAAACTTTTAGTCMAAGCATTTGCCAAAGCCATTTCATA	1080
Qy	364	GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu	383
Db	1081	GGAAATCTCTAAGGTTTCCGCTAAGGGAACCTCAATGATTGCCAAGACATCTCCCAAGCTC	1140
Qy	384	SerLysValIleSerSerLeuThrSerLysTSPValThrValGlyValGlyAlaValAl	403
Db	1141	TCGAAAGTCATCTGCTCTTACCACTAAATGGGTACGGTGGGGTGGGGTGTGTAATT	1200
Qy	404	AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGlnGln	423
Db	1201	GGCGGCGCTGCTCGGTAAAGGATTTATGCAATCGACTCTCGGAGATGACAAACAAAC	1260
Qy	424	ValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet	443
Db	1261	GTCCCTCAATTCAGAAAGAGTGGGAAAACTGCAAGCTGCGCGTGAATATGATTTCTATG	1320
Qy	444	PheThrGlnPheProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn	463
Db	1321	TTTCACTCAATTTTGGCAACGGGCAAGTAAATTTGGCTCAAAACCAAAACAGGCGAGTCTAAT	1380
Qy	464	GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaAlaTyrAlaAlaIle	483
Db	1381	GAATGACTCAAAAAGCTATCAAGAGTGGGCGCTCAAAATCCTTAAAGCGATGCCGCAATC	1440
Qy	484	SerLysAlaIleAlaGlyAla 490	
Db	1441	AGCGAGCCATTCGCTGGCGCA 1461	

RESULT 5	
AAT14616	
ID	AAT14616 standard; DNA; 5658 BP
XX	
AC	AAT14616;

XX		07-NOV-1996	(first entry)	C. pneumoniae antigen (polypeptide A) expression vector pCPN533 alpha
DT				
XX				
DE				pCNP533alpha; primer, assay; detection; antibody; diagnosis;
KW				Infection; Chlamydia pneumoniae; ds.
XX				Synthetic.
OS				
XX				
FH	Key	Location/Qualifiers	-	
FT	CDS	760..2847	/tag= "a"	
FT		/note= "STOP codon absent"		
XX				
PJ		WO9609320-A1.		
PD		28-MAR-1996.		
XX				
PF		20-SEP-1995:	95WO-JP01896.	
XX				
PR		28-APR-1995:	95JP-0106011.	
PR		20-SEP-1994:	94JP-0224711.	
PR		28-APR-1995:	95JP-0106006.	
PR		28-APR-1995:	95JP-0106008.	
PR		28-APR-1995:	95JP-0106009.	
PR		28-APR-1995:	95JP-0106010.	
XX				
PA		(HTB ) HITACHI CHEM CO LTD.		
XX				
PI	Izutsu H, Matsumoto A, Obara K;			
DR	WPI: 1996-188399/19.			
DR	P-PDB: AAR94579.			
XX				
PT	Recombinant Chlamydia pneumoniae antigen and antibodies to it -			
PT	used for detection and assay of C. pneumoniae e.g. in clinical			
PS	diagnosis			
Claim 13;	Pages 77-84; 128pp; Japanese.			
XX				
CC	The present sequence is an expression vector for the C. pneumoniae			
CC	polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was			
CC	cultured and genomic DNA extracted to prep. a lambda gt11 DNA			
CC	library. The library was then screened with an anti-YK41			
CC	monoclonal antibody (mAb), which was prepd. by fusing spleen cells			
CC	from a mouse infected with YK41 with myeloma P3/NSI/1-Tg4-1 to			
CC	produce a MAB expressing hybridoma. The DNA obt. was then fused			
CC	with the expression vector PAD431 to give pcPN533alpha. The			
CC	plasmid was used to transform an E. coli host, which was cultured			
CC	to give the antigenic polypeptide, polypeptide A. Polypeptide A			
CC	and primers and probes derived from its DNA can be used in assays			
CC	for the detection of polypeptide A antibodies and DNA,			
CC	respectively, useful in the diagnosis of C. pneumoniae infection.			
XX				
SO	Sequence 5658 BP; 1446 A; 1481 C; 1424 G; 1307 T; 0 other;			
Alignment Scores:				
Pred. No.:	9.8e-86	Length:	5658	
Score:	1282.00	Matches:	315	
Percent Similarity:	62.59%	Conservative:	73	
Best Local Similarity:	56.65%	Mismatches:	35	
Query Match:	54.12%	Indels:	135	
DB:	17	Gaps:	10	
US-09-889-314-2 (1-496) x AAAT14616 (1-5658)				
OY	1 AspTrrAsnMetSerLileSerrSerSerglyProaspAsglnLyAsnIlleMeSer 20			
Ddb	 751 GATTACAACAGTCGTCAITTTCTCATCTTGCGACCCTGCATCAAAAATAATGATGTC 810			
OY	21 GluValLeuThrSerTrhPrroGlncglYvalProGlncglnaSplysLeuSerglYAsgcgu 40			

Db 1384 GGTGCCGCAACGACGCTACCAACAAATTACAGTTCAAGCTGTGTCTCCAAAGCGGTGAAA 1443  
 QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaValLys 343  
 Db 1444 CAAGCTGTTATCAGCTGCTGACAGCAAGGATACCGCGGCTATTAAGCGGCTGTCAAA 1503  
 QY 344 SerGlyIleLysAlaIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
 Db 1504 TCTGGATATAAAGCATTTATCAAAACTTATAGTCAAAAGCATTTGCCAAAGCATTTCTAAA 1563  
 QY 364 GlyIleSerLysValaIleAlaLysGlyThrGlnMetIleAlaLysAsnPhenProLysLeu 383  
 Db 1564 GGAATCTTAAGGTTTTCGCTAAGGAACTCAATGATGCGAAGACTTCCCAAGCTC 1623  
 QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403  
 Db 1624 TCCAAAGTCACTCTGCTCTTACAGTAATGGTTCACGGTTGGGGTGGGTGTAGTT 1683  
 QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnLeuGlnGlnAsn 423  
 Db 1684 GCGGCGCTCTCTCGGTAAGGATTTATGCAATGATGCTTCGGAGATGCAACAAAC 1743  
 QY 424 ValAlaIleGlnGlnLysGlyValGlyLysLeuGlnAlaIleAlaAspMetIleSerMet 443  
 Db 1744 GTGCTCAATTTTCAGAAAGAGTCGAAACTGCAAGCTCGGCTGTATGATTTCTATG 1803  
 QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSerAsn 463  
 Db 1804 TTCTCTCAATTTTGGCAACAGGCAAGTAATTTGCTCAAAACAAACAGCGAGTCTAAT 1863  
 QY 464 GluMetThrGlnLysAlaIleThrLysLeuGlyAlaGlnIleLeuLysAlaIleAlaIle 483  
 Db 1864 GAAATGACTCAAAAGCTACCAAGCTGGGCGTCAATCTTAAAGCTATGCGGCATC 1923  
 QY 484 SerGlyAlaIleAlaGlyAla 490  
 Db 1924 AGCGAGCGCATCGCTGGCGCA 1944

RESULT 4  
 AAT14612  
 ID AAT14612 standard; DNA: 1464 BP.  
 AC AAT14612;  
 XX  
 DT 07-NOV-1996 (first entry)  
 XX  
 DE Chlamydia pneumoniae polypeptide antigen (polypeptide A) DNA.  
 XX  
 KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;  
 KW PCN533alpha; primer; assay; detection; antibody; diagnosis;  
 KW Infection; ds.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 FH Key 1..1464  
 FT CDS Location/Qualifiers  
 FT 1..777 /\*tag= a  
 FT misc\_feature /note= "STOP codon absent"  
 FT /\*tag= b  
 FT /\*note= "claimed 5'-end fragment"

PR 28-APR-1995; 95JP-0106010.  
 XX  
 PA (HITB ) HITACHI CHEM CO LTD.  
 XX  
 PI Izutsu H, Matsumoto A, Obara K;  
 XX  
 DR WPI: 1996-188399/19.  
 DR P-PSDB: AAR94579.  
 XX  
 PT Recombinant Chlamydia pneumoniae antigen and antibodies to it  
 PT used for detection and assay of C. pneumoniae e.g. In clinical  
 PT diagnosis  
 XX  
 PS Claim 9; pages 60-64; 128pp; Japanese.

PR 28-APR-1995; 95JP-0106010.  
 XX  
 PA (HITB ) HITACHI CHEM CO LTD.  
 XX  
 PI Izutsu H, Matsumoto A, Obara K;  
 XX  
 DR WPI: 1996-188399/19.  
 DR P-PSDB: AAR94579.  
 XX  
 PT Recombinant Chlamydia pneumoniae antigen and antibodies to it  
 PT used for detection and assay of C. pneumoniae e.g. In clinical  
 PT diagnosis  
 XX  
 PS Claim 9; pages 60-64; 128pp; Japanese.

The present sequence encodes the C. pneumoniae polypeptide antigen,  
 polypeptide A. C. pneumoniae strain YK41 was cultured and genomic  
 DNA extracted to prep. a lambda gt11 DNA library. The library was  
 then screened with an anti-YK41 monoclonal antibody (MAb), which  
 was prep. by fusing spleen cells from a mouse infected with YK41  
 with myeloma F3/NS1/-Ag4-1 to produce a MAb expressing hybridoma.  
 The DNA obd. was then fused with the expression vector pAD431 to  
 give pcPN533alpha. The plasmid was used to transform an E. coli  
 host, which was cultured to give the antigenic polypeptide,  
 polypeptide A. Polypeptide A and primers and probes derived from  
 its DNA can be used in assays for the detection of polypeptide A  
 antibodies and DNA, respectively, useful in the diagnosis of  
 C. pneumoniae infection.

Sequence 1464 BP; 470 A; 333 C; 346 G; 315 T; 0 other:  
 SQ  
 Alignment Scores:  
 Pred. No.: 3,656-162 Length: 1464  
 Score: 2300.00 Matches: 485  
 Percent Similarity: 99.598  
 Best Local Similarity: 99.598  
 Query Match: 97.098  
 DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AAT14612 (1-1464)  
 QY 4 MetSerLleSerSerSerSerGlyProAspGlnLysAsnIleMetSerGlnValLeu 23  
 Db 1 ATGCTATTTCAATCTCTTCTCAGACCTCGACAAATCAAAATATCTCAAGTTCTG 60  
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 Db 61 ACATCGACACCCAGGCGCGCCCAACAAAGATTAAGCTGTCTGCAACGAAGCAAGCA 120  
 QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGluMetGlnSerAspAlaThrIleAlaGly 63  
 Db 121 ATACAGCAACACGCTCAGGTTAAACACAGATGGAAGCATGCGACTATTGCTGGT 180  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGln 83  
 Db 181 GCTTCTGAAAAGACAAACTCTCTCGACTACCAAAACAAACAAACAGCTCCACAAAGGA 240  
 QY 84 ValAlaIleGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
 Db 241 GTTCTGCTGGGAAGAATCTCTGGAAGTCAAAAGCAGGTGCTGATCTGGAGTATCA 300  
 QY 104 GlyAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 Db 301 GGAGCGGCTCTACTACAGCATCAAAATCTGCAAAATTTGCTAAGCAGACTCTATT 360  
 QY 124 GlnGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 Db 361 GAAGAGCGGACCAAAAGATGAGTCTTACCTTAAGTCACTTCAAAACCTCAGCGCGG 420  
 QY 144 GlnMetLysGlyValGluAlaValAlaValAlaIleLysSerGlyLysSerSerGlySer 163  
 Db 421 CAATGAAAGAGTCAAGCGGTTGTGCTGCTCCCTCTCAGGGAAGTTCGGGTTCC 480

```
QY 461 GUSERASGLUMETThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480
DB 926540 GAGCTATATGAATGATCATCAAAAGCTACCAAGCTGGCGCTCAAAATCTTTAAAGCGTAT 926481
QY 461 AlaAlaIleSerGlyAlaIleAlaGlnLysThrAsnAsnIle 496
DB 926480 GCCGCAATCAGCGGAGCCATCTGGCGCAGCATAAACCAATAATTTT 926433

RESULT 3
AAT14618
ID AAT14618 standard; DNA; 1947 BP.
XX
AC AAT14618;
XX
DT 11-NOV-1996 (first entry)
DE DHFR/C. pneumoniae antigen fusion protein (polypeptide B) DNA.
XX
KW Polypeptide antigen; polypeptide B; strain YK41; plasmid; probe;
KW PCPN533T; primer; assay; detection; antibody; diagnosis;
KW infection; fusion protein; dihydrofolate reductase; DHFR;
KW Chlamydia pneumoniae; ds.
XX
OS Synthetic.
XX
FH Key 1.1947 Location/Qualifiers
FT CDS /*tag= a
FT /note= "STOP codon absent"
XX
PN MO9609320-A1.
XX
PD 28-MAR-1996.
XX
PE 20-SEP-1995; 95WO-JP01896.
XX
PR 28-APR-1995; 95JP-0106011.
PR 20-SEP-1994; 94JP-0224711.
PR 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX
PA (HITB) HITACHI CHEM CO LTD.
XX
PI Izutsu H, Matsumoto A, Obara K;
XX
DR WPI; 1996-188399/19.
XX
DR P-PSDB; AAR94584.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it -
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX
XX Claim 25; Pages 97-103; 128pp; Japanese.
XX
XX The present sequence encodes the dihydrofolate reductase (DHFR)/
XX C. pneumoniae polypeptide antigen fusion protein, polypeptide B.
XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted
XX to prep. a lambda gt11 DNA library. The library was then screened
XX with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
XX fusing spleen cells from a mouse infected with YK41 with myeloma
XX P3/NS1/1-2g4-1 to produce a Mab expressing hydridoma. The DNA
XX obdd. was then fused with DHFR DNA and the expression vector
XX pAD431 to give pCPN533T. The plasmid was used to transform an
XX E. coli host, which was cultured to give the antigenic polypeptide
XX fusion protein, polypeptide B. Polypeptide B and primers and
XX probes derived from its DNA can be used in assays for the
XX detection of antigenic polypeptide antibodies and DNA,
XX respectively, useful in the diagnosis of C. pneumoniae infection.
XX
XX Sequence 1947 BP; 591 A; 452 C; 483 G; 421 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 9,24e-163 Length: 1947
Score: 2310.00 Matches: 485
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.59% Mismatches: 0
Query Match: 97.51% Indels: 0
DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AAT14618 (1-1947)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23
DB 484 ATGTCTATTTTCATCTTCTTCAGAGACCTGACATCAAAAAATATCATGTCTCAAGTCTCG 543
QY 24 ThrSerThrProGlnGlyValProGlnAsnAspLysLeuSerGlyAsnGlnThrLysGln 43
DB 544 ACATCGACACCCCGAGCGCTGCCCAACAGATAGCTGTGGCAACGAAACGAAACGAA 603
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
DB 604 ATACAGCAACACGCTCAGGCTAAACACTGAGATGGAAGCGATGCCCTATTGCTGGT 663
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
DB 664 GCTTCGGAAGAGACAAACCTCTCGACTACAAAACGAAACGAGCTCCACACAGCGGA 723
QY 84 ValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 724 GTTGCTGTGGGAAAGAAATCTCAGAAAGTCAAAAGGCGGTGTGATCTGAGATATCA 783
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 784 GGAGCGGCTGCTACTACACATCAAAATCTGACACAAAATTTGCTTAGACCTCTATT 843
QY 124 GluGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuAlaAla 143
DB 844 GAAGAGCGGAGCAAAAGTATGAGTCTACCTTACAGTCACTCAAAAGCCTCAGTGGCGG 903
QY 144 GlnMetLysGlnValGlnAlaValValValAlaAlaLeuSerGlyLysSerSerGlySer 163
DB 904 CAATGAAAGAAAGTCAACCGGTTGTGTGCTGCTCAGGGAAGGTTCCGGGTTCC 963
QY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
DB 964 GCAAAATTTGAAACACCTACCTCCCCAAGCCCGGGGTACACCAAGATCAAGGTTATTC 1023
QY 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
DB 1024 GAATCGGACCTCGCGCTGCTTAAGCAATTCAGACATTGGAGAAAGCCAAATCTGCC 1083
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGlu 223
DB 1084 TTATCTTAACCTATGCAAGTCAACAGCAACGACCAACCAATTAATCTAGTCTAGAA 1143
QY 224 LysGlnAlaIleLysIleAspLysGluArgGlnGluThrGlnGlnMetLysAlaIleGlu 243
DB 1144 AAGCAAGCGATTAATAATCATTAAGACAGAAAGATTAACCAAGATGAAGCTGCGGAA 1203
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 1204 CAGAAGTCAAAAGATCTCAAGAGCAACATGATGCTCAATATGATGATCGGGTT 1263
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
DB 1264 TCTGTGTCATTAACAGTTTCTATTGTGTGCTGCTATTTTACATGCGGACCTGGACAC 1323
QY 284 AlaGlyLeuAlaAlaGlyAlaValAlaGlyAlaAlaAlaAlaGlyValAlaIleAlaGlyAla 303
DB 1324 GCTGCACTGCTGCGGAGCTGCTGTAGGTGCAAGCGGACGCTGGAAGTCCACAGAGAGCT 1383
QY 304 AlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323
```



XX AAX91990;  
 AC 13-SEP-1999 (first entry)  
 DT  
 XX Nucleotide sequence of the complete genome of Chlamydia pneumoniae.  
 DE  
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KM vaccine; neutralising epitope; ss.  
 XX Chlamydia pneumoniae.  
 OS  
 XX W09927105-A2.  
 PN  
 XX 03-JUN-1999.  
 PD  
 XX 20-NOV-1998; 98MO-IB01890.  
 PE  
 XX 04-NOV-1998; 98OS-0107078.  
 PR 21-NOV-1997; 97FR-0014673.  
 PA (GEST ) GENSET.  
 XX  
 PI Griffiths R;  
 XX  
 DR WPI; 1999-357842/30.  
 XX  
 PT Genome sequence of Chlamydia pneumoniae  
 PS  
 XX Claim 1; Page 291-611; 1912pp: English.  
 CC The present sequence represents the complete genome of Chlamydia pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes respiratory disease such as pneumonia and bronchitis, and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in immunogenic compositions as vaccines. CC Vectors containing C. pneumoniae nucleotide sequences can also be used as immunogenic compositions, especially where the vector directs CC the expression of a neutralising epitope of C. pneumoniae.  
 XX  
 SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;  
 Alignment Scores:  
 Pred. No.: 1.02e-162 Length: 1230025  
 Score: 2354.00 Matches: 493  
 Percent Similarity: 99.40% Conservative: 0  
 Best Local Similarity: 99.40% Mismatches: 3  
 Query Match: 99.37% Indels: 0  
 Gaps: 0  
 DB:  
 US-09-889-314-2 (1-496) x AAX91990 (1-1230025)  
 OY 1 AsptRrasmMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20  
 DB 927920 GATGAAACATGCTATTTCATCTCTTCAGAGACTGACATCAAAAAATATCTGCT 927861  
 OY 21 GlnValIleuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40  
 DB 927860 CAAAGTTCGACATCGACACCCAGGCGTCCCAACAAAGATTAAGCTGCTGCAACGAA 927801  
 OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspLysThr 60  
 DB 927800 ACGAAGCAAAATACACAAACAGTCAAGGTAATAAACACATGAGATGAAAGCCATCCACT 927741  
 OY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80  
 DB 927740 ATTGCTGGTCTTCTGGAAGAAACAAACTCTCTGACGTACAAAAACGAAACAGCTCA 927681  
 OY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100

162 (2) m 2003

DB 927680 CAACAGGAGATTGCTGCGGGAAGATCTCGAAAGTCAAAAGCAGCGTCTGATACT 927621  
 OY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
 DB 927620 GGAGTATCAGAGAGCGGCTCTACTACACATCAAAATGCAACAAAAATTTCTATGCG 927561  
 OY 121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGlnLysLeuSerLeu 140  
 DB 927560 ACCCTATTGAAAGAGCGGACCAAAAGTATGAGACTCTACTTGAAGCATTTCAAGCCCTC 927501  
 OY 141 SerAlaAlaGlnMetLysGluValAlaValAlaValAlaAlaLeuSerGlyLysSer 160  
 DB 927500 AGTCCGCGCAAAATGAAAGAGTCGAAGCGGTGTGCTGCTCCCTCCAGGGAAGT 927441  
 OY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
 DB 927440 TCGGGTCCCGCAAAATGGAACACCTGAGCTCCCAACCGCGGGGTGACCCAGATCA 927381  
 OY 181 GluValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysAlaThr 200  
 DB 927380 GAGGTTATCGAAATTCGACCTCGCGCTTGCCTTAAGCAATTCAGACATGGGAGAGCCACA 927321  
 OY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
 DB 927320 AAATCTGCTTATCTAATCTATGCAAGTACACAAAGCACCAACCAACCAATTAACGA 927261  
 OY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlnLysGlnLysGlnMetLys 240  
 DB 927260 GGCTTACAAAAGCAAGGATTAATAATGATTAAGGACGAAACAAATACCAAGAGATGAAG 927201  
 OY 241 AlaAlaGlnLysSerLysAspLysGlnGlyThrMetAspThrValAsnThrValMet 260  
 DB 927200 GCTCCGCAACAGAAAGTCAAAAGATCTCGAAGGAACAATGATCTGCTGATG 927141  
 OY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaIlePheThrLysGly 280  
 DB 927140 ATCCGGTTCCTGTTGTCATTAACGATTAATTCATGTTCTCTCTATTTTACATGGGA 927081  
 OY 281 AlaGlyLeuAlaGlyLeuAlaAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAla 300  
 DB 927080 GCTGAGCTCGTGTGACTCGCTGCGGAGCTGCTGTAGTGTGACGCGGAGCTGGAGGTGCA 927021  
 OY 301 AlaGlyAlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValGln 320  
 DB 927020 GCAGAGACTGCTGCCGCAACACGCTGACCAACAAATTCAGTTACAGTGTGTCCAA 926961  
 OY 321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340  
 DB 926960 GCGGTGAACAAAGCTGTTATACAGCTGTGACAGACAGCATCACCGCGCTATTAAGCG 926901  
 OY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
 DB 926900 GCTGCAAAATCTGGAATTAAGCATTAATCAAACTTATGCAAAACGATTCGCAAAAGCC 926841  
 OY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380  
 DB 926840 ATTCTTAAGGAATCTTAAGGTTTCCGTAAAGCACTCAAAAGATGCGAAGACATTC 926781  
 OY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGly 400  
 DB 926780 CCCAAGCTCTGAAAGTATCTCGCTCTTACAGTAATAGGTCAGCGTTGGGGTGGG 926721  
 OY 401 ValValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMet 420  
 DB 926720 GTTGTAGTTGGCGCGCTCTCTCGTAAAGGATTAATGCAAAAGTCAGACTCTCGAAGATG 926661  
 OY 421 GlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMet 440  
 DB 926660 CAACAAAACGTCGCAATTCAGAAAGAGTCGGAAGAACTGCAAGCTCGCGCTGATATG 926601  
 OY 441 IleSerMetPheThrGlnPheThrProGlnLysAlaSerLysIleAlaSerLysGlnThrGly 460  
 DB 926600 ATTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCTCAAAACAAACAGCG 926541



```

XX  MO200046359-A2.
PN  10-AUG-2000.
XX  28-JAN-2000; 2000WO-GB00237.
XX  05-FEB-1999; 99GB-0002555.
XX  (NEUT-) NEUTEC PHARMA PLC.
PA  Burnie JP, Matthews RC;
PI  WPI: 2000-543485/49.
XX  P-PSDB: AAB08322.
DR  New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX  preventing and treating C. pneumoniae infection and atherosclerosis,
PT  including coronary atherosclerosis -
XX  Claim 3; Page 27-29; 35pp; English.
XX  The present sequence encodes a Chlamydia pneumoniae protein. The
CC  protein, immunogenic fragments of it, nucleotide sequences encoding
CC  it, or inhibitor specific against it are used to manufacturing
CC  a medicament for the treatment of infection due to C. pneumoniae.
CC  An antibody specific against the protein can diagnose a C. pneumoniae
CC  infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC  including coronary atherosclerosis, caused by C. pneumoniae can also
CC  be prevented or treated.
XX  Sequence 1491 BP; 485 A; 335 C; 348 G; 323 T; 0 other:
XX
Alignment Scores:
Pred. No.: 2,75e-167 Length: 1491
Score: 2369.00 Matches: 496
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-889-314-2 (1-496) x AAA63621 (1-1491)
OY  1 AsptRrAmSeSerThrSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
DB  1 GATACAAACATGTTCTATTTCATCTTCTCAGACCTGACAAATCAAAAATATTCATGTT 60
OY  21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40
DB  61 CAAGTTCTGACATCGACACCCAGGGCGTGGCCCAACAAGATTAAGCTGTGCAACGAA 120
OY  41 ThrLysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThr 60
DB  121 ACGAAGCAAAATACACCAAAACGTCAGGGTAAACACATGATGGAAGGCATGGCACT 180
OY  61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80
DB  181 ATTGCTGTGCTTCTTGGAAGAACAAACTCTCTGACTACAAAACACAAACAGCTCCA 240
OY  81 GlnGlnGlyAlaIleAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100
DB  241 CAACAGGGAGGTGCTGCTGGGGAAGAAATCTCAGAAAGCAAAAGCAGCTGATACT 300
OY  101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB  301 GGAGATATCAGGCGCGCTCTACTACAGCATCAAAATACTGCAACAAATTCCTATGCG 360
OY  121 ThrSerIleGlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
DB  361 ACCTCTATTGAAGAGCGGCAAAAGATGAGATCTACCTTAGAGTCACCTTAAGACCTC 420
OY  141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaIleLeuSerGlyLysSer 160
DB  141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaIleLeuSerGlyLysSer 160

```

```

DB  421 AGTCCCGCAAAATGAAGATCGAAGCGGTTGTTGCTGCCCTCTCAGGCAAAAGT 480
OY  161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSer 180
DB  481 TCGGGTCCGCAAAATGGAACACCTAGCTCCCAAGCCGGGGGTGACACCAAGATCA 540
OY  181 GlnValIleGlnIleGlyLysLeuAlaLeuAlaIleGlnThrLeuGlnGlyAlaThr 200
DB  541 CAGCTTATCGAAATGCGACTCCGGCTTGCTAAAGCAATTCACACTTGGGAAACCA 600
OY  201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB  601 AAATCTGCCCTTATCTAATCATCTCAAGTACACAAAGCAAGCAGCAACAAATTAAC 660
OY  221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnValGlnGlnGlnGlnGlnGln 720
DB  661 GGTCTAGAAAGCAAGCGATTAATAATCGATTAAGAAGCAAGAAATACCAAGAGATGAG 720
OY  241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
DB  721 GCTGCCGAACGAAAGTCTAAAGATCTCGAAGAACAAATGATACTGTCATCTGTGATG 780
OY  261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280
DB  781 ATCGGGGTTCTGTGGCATCTACAGTATCTAATGTTGCTGCTATTTTACATGCGGA 840
OY  281 AlaGlyLeuAlaGlyLysLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaIleGlyAla 300
DB  841 GCTGACCTCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
OY  301 AlaGlyAlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValGln 320
DB  901 GCAGAGCTGCTGCTGCCGAACACCGGTAGCAACAAATTAACAGTTCACTGTGCGCA 960
OY  321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
DB  961 GCGGTGAACCAAGCTGTATCTACAGCTGTCAGCAAGCATCACGCGCTTAAAGCG 1020
OY  341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
DB  1021 GCTGTCAAAATGGAATTAACCAATTAATCAAAACCTTAAGTCAAAAGCGATTCGCA 1080
OY  361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
DB  1081 ATTCTTAAGGAATCTCTAAGCTTTTCCTTAAGGCAACCAATTAATTCGAAGAATCT 1140
OY  381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyAla 400
DB  1141 CCCAAGCTCTCGAAAGTCATCTCGTCTTACCAAGTAATGGGTCCAGGTGGGTTGGG 1200
OY  401 ValValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGln 420
DB  1201 GTTGTAGTTCGCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
OY  421 GlnGlnAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMet 440
DB  1261 CAACAAAGCTGCTCAATTCAGAAAGAAAGTCGGAAGAACTCGAGGCTCGGCTGATAG 1320
OY  441 IleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460
DB  1321 ATTTCTATGTCATCAATTAATTTGGCAACAGGCAAGTAATTAATTTGCCCAAAACAG 1380
OY  461 GlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThr 480
DB  1381 GAGCTTAATGAATGATCACTCAAAAGCTACCAAGCTGCGGCTCAATCTTTAAAGGTAT 1440
OY  481 AlaAlaIleSerGlyAlaIleAlaIleAlaIleLysThrAsnAsnPhe 496
DB  1441 GCCGCAATCAGGGAAGCATTCGCTGGCGCAATTAACCAATTAATTTT 1488

```

RESULT 2  
 AAX91990/C  
 ID AAX91990 standard; DNA; 1230025 Bp.

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## SUMMARIES

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 27, 2003, 14:05:52; Search time 219 Seconds  
(without alignments)  
5100.416 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 2369  
1 DTNMSISSSGPDNOKNIMS.....LKAYAAISGALGAHKTNFF 496

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O=cgq2.1/USPTO.spool/US09889314/rnat\_24012003\_144539\_10097/app\_query.fasta\_1.647  
-DB=N-Geneseq.101002 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOPCCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=45  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09889314 -CGCN\_1.1.125 -runat\_24012003\_144539\_10097 -NCPV=6 -ICPV=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	2369	100.0	1491	21 AAA63621	DNA encoding a Chl
2	2354	99.4	1230025	20 AAX91990	Nucleotide sequenc
3	2310	97.5	19447	17 AAT14618	DHFR/C. pneumoniae
4	2300	97.1	1164	17 AAT14616	Chlamydia pneumoni
5	1282	54.1	5658	17 AAT14616	C. pneumoniae anti
6	1266	53.4	5439	17 AAT13893	C. pneumoniae anti
7	1262	53.3	1048	17 AAT14622	C. pneumoniae poly
8	1262	53.3	1048	18 AAT59311	DNA encoding C. pn
9	1245	52.6	813	17 AAT14613	C. pneumoniae poly
10	1245	52.6	1296	17 AAT14619	DHFR/C. pneumoniae
11	1195.5	50.5	1038602	20 AA201425	Complete genome se
12	213.5	9.0	32768	19 AAV52204	Streptococcus pneu
13	199	8.4	4483	19 AAV52496	Streptococcus pneu
14	196	8.3	3633	21 AAA05548	Streptococcus pneu
15	193	8.1	1702	19 AAV52500	Streptococcus pneu
16	192	8.1	2550	19 AAV52497	DNA sequence that
17	191.5	8.1	15620	19 AAV03553	Streptococcus pneu
18	188	7.9	1436	19 AAV52498	Streptococcus pneu
19	187	7.9	8201	21 AAB88864	Human dentin stalo
20	187	7.9	4638	23 ABL17521	Drosophila melanog
21	186.5	7.9	6771	23 ABL04818	Shigella flexneri
22	186.5	7.9	18869	23 ABL04818	Botrellia variable
23	186.5	7.9	7766	18 AAT85043	Shigella flexneri
24	186	7.8	1743	18 AAT85939	N. meningitidis pa
25	185	7.7	49914	21 AAA81471	Neisseria meningit
26	182.5	7.7	349980	21 AAF21607	Lactobacillus rham
27	182.5	7.7	1437668	21 AAA81490	Streptococcus poly
28	182.5	7.6	4645	24 AAD31861	Drosophila melanog
29	180	7.6	2910	24 AAB68049	Human S3-12 homolo
30	179	7.5	3750	22 ABA09498	Streptococcus pneu
31	178	7.4	869	24 AAV52502	Human secreted pro
32	174.5	7.3	6108	24 ABL39666	Drosophila melanog
33	172.5	7.2	10910	24 ABL03131	F. necrophorum leu
34	171.5	7.1	11130	24 AAT67647	Staphylococcus aur
35	168.5	7.1	941	19 AAV52396	F. necrophorum leu
36	168.5	7.1	9726	23 AAS54978	Streptococcus pneu
37	167.5	7.1	9726	24 AAT67640	Streptococcus pneu
38	167.5	7.0	3744	19 AAV52501	Drosophila melanog
39	167	7.0	941	19 AAV52396	Streptococcus pneu
40	166.5	7.0	2708	23 ABL19629	Streptococcus pneu
41	166	7.0	780	19 AAV52503	Listeria innocua C
42	165.5	7.0	1163020	24 ABO67197	Listeria innocua D
43	165.5	7.0	3011208	24 ABO69245	Nucleotide sequenc
44	165.5	7.0	3725	22 AAT78245	
45	165	7.0	3725	22 AAT78245	

## ALIGNMENTS

RESULT 1  
ID AAA63621 standard: DNA: 1491 BP.

AAA63621:

04-DEC-2000 (first entry)

DNA encoding a Chlamydia pneumoniae protein.

Chlamydia pneumoniae protein; infection: arteriosclerosis;

coronary atherosclerosis; ss.

Chlamydia pneumoniae.

Key Location/Qualifiers

FT CDS 1..1491 /tag= a

```
OY 400 GlyValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlu 419
      ::::: ||||| :::: |||
Db 3202 AATATCATTTCAAAA-----AACGGTATTAACACCGTACTGTATAAAGGC 3246
      ::::: ||||| :::: |||
OY 420 MetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAsp 439
      ::::: ||||| :::: |||
Db 3247 GTTAAATTTGATGTGAATACATTCACCGCGGTATA-----GCAAGCGTAGAT 3294
      ::::: ||||| :::: |||
OY 440 MetIleSerMetPheThrGlnPheThrPglGlnAlaSerLysIleAlaSerLysGlnThr 459
      ::::: ||||| :::: |||
Db 3295 GAAGTAATT-----GAAGCGAAACGCATCCTTGAGAGGTAATAA 3333
      ::::: ||||| :::: |||
OY 460 GlyGlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAla 479
      ::::: ||||| :::: |||
Db 3334 GATTATCTGATGAAGAAGAAGACCGTTAGCTAAACTTGGCGTAAGCGCTGTACGT--- 3390
      ::::: ||||| :::: |||
OY 480 TyrAlaAlaIleSerGlyAlaIle 487
      ::::: ||||| :::: |||
Db 3391 TTTGCTGAGCCAAATATATGCCATT 3414
      ::::: ||||| :::: |||
```

Search completed: January 27, 2003, 16:22:19  
Job time : 3137 secs

```

OY 341 a-----VallySerGlylleYsAlaPheIleYsThrLeuVallyAlaI 357
|          |||          |||          |||          |||
Db 2943 CTTACCCCCCAAGCCTCTCGCGCCCCCGCCCTGCTGTGCAGAGCTGCAAGCAG 3002

OY 357 eAlaYsAlaIleSerlySlyIleSerlySValPheAlaLyGlyThrGlnMetIleAl 377
|          |||          |||          |||          |||
Db 3003 GGCAGAGCAAGATTCCACTGCTGTGCAGGCGTC-----CGAGGAAGCCAA--GCCCA 3053

OY 377 alyAsnProIlySerlySerlySValIleSerSerIleUThrSerlyTrpValThrVa 397
|          |||          |||          |||          |||
Db 3054 GCCTGACAGCCCAAGCCTGAGCTTGCCCTCATTTGCTGCCAGCAAGAGCTTCCTCCAGCC 3113

OY 397 lGlyValGlyValValAlaAla-----ProAlaLeuGlyLySgIlyIleMe 413
|          |||          |||          |||          |||
Db 3114 AGGTGGAGAGATGGTGGCAGCTGCAGGAAGCCTCAGTCCCAACGATTCAGGACCAAGCCTTC 3173

OY 413 tGlnMetGlnLeuSerGlnMetGlnGlnAn-----ValAlaGlnPheGlnIly 429
|          |||          |||          |||          |||
Db 3174 AGCCATGCAGCTGAGTCAAGTGTGCCAGGAACCTGGGCAAGCCGCGCTGCTGAATCCGGAC 3233

OY 429 sGluValGlyLySLeuGlnAlaAlaAlaAspMetIleSerMet 443
|          |||          |||          |||          |||
Db 3234 GCGTGGCCCAAGAGCCTCAGAGAGCATGTGAGACCTTTGGAGATG 3276

RESULT 45
US-09-206-942-64
; Sequence 64, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; FILE REFERENCE: 1038-861 MIS:jd
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 64
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; US-09-206-942-64

Alignment Scores:
Pred. No.: 9.75e-05 Length: 3543
Score: 146.50 Matches: 114
Percent Similarity: 36.55% Conservative: 79
Best Local Similarity: 21.59% Mismatches: 228
Query Match: 6.18% Indels: 107
DB: 4 Gaps: 23

US-09-889-314-2 (1-496) x US-09-206-942-64 (1-3543)

OY 7 SerSerSerSerGlyProAspAsnGlnIlySAsnIleMetSerGlnValLeuThrSerThr 26
|          |||          |||          |||          |||
Db 2011 AACACTGAAGTGGCAGAGACAAACAATACCGCT-----TTAACTATTACT 2055

OY 27 ProGlnGlyValProGlnGlnAspIlySLeuSerGlyAsnGlnIlyThrLySglnIleGln 46
|          |||          |||          |||          |||
Db 2056 GCAAAAGATGTAGAAATGAACAACATATTACTTCAACAAACAGTGAAGCTCTCGCG 2115

OY 47 ThrArgGlnGlyLyAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyValAserGly 66
|          |||          |||          |||          |||
Db 2116 GCAAAATGAGGATTAACCT-----AAACAGGTACACCATTAATCCACCGCGGT 2169

OY 67 LysAspLySThrSerSerThrThrIlySThrGlnIlyThrAlaProGlnGlnIlyValAlaAla 86
|          |||          |||          |||          |||
Db 2170 AACGGGAGATTAACCGCT-----CATACAGCAGTATCCAAAGCGCAATTGAG 2217

```

```

OY 87 GlyLySgluSerSerGluSerGlnIlySAlaGlyValAspThr-----GlyVal 102
|          |||          |||          |||          |||
Db 2218 TCCAGCTCGCTCTGTGTGCAATTTGTGGCAGCGCGGATACTCTTGCTAGTAAATTT 2277

OY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrIlySAlaMetGlnThrSer 122
|          |||          |||          |||          |||
Db 2278 TCAGGCAAGCGCGTTACTCTTACTGCAATTAACCGT-----GCATTAACCACTTTG 2328

OY 123 lIleGlnAlaIleSerlySerMetGluSerThrLeuGlnSerLeuGlnSerIleuSerAla 142
|          |||          |||          |||          |||
Db 2329 GCAGGCTTACAATTAAGAAACGAGAGTAAACCACTTCAAGTCAATCAAGTAAATATC 2388

OY 143 AlaGlnMetLySgluValGlnAlaValValAlaAlaLeuSerGly--LysSerSer 161
|          |||          |||          |||          |||
Db 2389 GCGGTAATAATTTCCGCGCAAGACAGTAACGTTAAAGCAACTAATAGTTAACCAACCCA 2448

OY 162 GlySerAlaLySLeuGlnIlyThrProGlnLeuProLySProGlyValThrProArgSerGlu 181
|          |||          |||          |||          |||
Db 2449 GCAGACTCAAAATTTGAGGAGACTGAA--GGCGAGCTAATGTACACAGCAAAACAAAGC 2505

OY 182 ValIle-----GluIleGlyLeuAlaLeuAlaLySAlaIleGlnThrLeuGlyGlnAla 199
|          |||          |||          |||          |||
Db 2506 ATAAATGGCGGTACAATTTCTGTGGCACAGTAAGATTACCGCGCAGAGGTTTAACC 2565

OY 200 ThrLySerAlaLeuSerAsnTyraIleSerThrGlnAla-----GlnAlaAspGlnThr 217
|          |||          |||          |||          |||
Db 2566 ACCCAAGCAGGC---TCTACGATTACTGGAACCGAGACGCTGACCACTTCAAGCCAAATCA 2622

OY 218 AsnLySLeuGly-----LeuGlnLySAlaIleIlys----- 228
|          |||          |||          |||          |||
Db 2623 GGTAAATATCGCGGATGATTTCTGTGGCAAGTGAAGATTAGCCCAACCAAGATTTA 2682

OY 229 lIleAspLySgluValGlnIlyThrGlnMetLySAlaAlaGlnIlySLeuSerLySAsp 248
|          |||          |||          |||          |||
Db 2683 ATTAATCAATCCGCTTCAGAGATT--AAAGCAAGCGCGGCGAGGTGAATGAATGAAGT 2739

OY 249 LeuGlnIlyThrMetAsp-----ThrValAsnThrValMetIleAlaValSerVal 265
|          |||          |||          |||          |||
Db 2740 GCAACAGATGCATATGACGAGTACGATTTCCGTAATATCGTAATGTACAGCAAAATATAC 2799

OY 266 AlaIleThrValIleSerIleValAlaAlaIlePheThrLySgIlyAlaGlyLeuAlaGly 285
|          |||          |||          |||          |||
Db 2800 GCGCATTTAACTGTGAAGATGCCGCAAAATATGATCGACAGCA----- 2844

OY 286 LeuAlaIleGlyAlaAlaValaGlyAlaAlaIleGlyAlaAlaGlyAlaAlaAlaAla 305
|          |||          |||          |||          |||
Db 2845 -----GGAGCGCGGACC 2856

OY 306 AlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaValGlnAlaValLySglnAla 325
|          |||          |||          |||          |||
Db 2857 CTAACGTCAACATCGGCAAAATTAACCACTTAAGGCT-----AGTTCAAGC 2901

OY 326 ValIleThrAlaValaValGlnAlaIleThrAlaAlaIleLySAlaAlaValLySergly 345
|          |||          |||          |||          |||
Db 2902 ATTAATTCAGCTAATTAACCAAGTAAACCTTCAAGTAAAGATGTGATTTGGGGGAAT 2961

OY 346 lIleYsAlaPheIleIlySThrLeuVallySAlaIleAlaIlySAlaIleSerLySgIlyIle 365
|          |||          |||          |||          |||
Db 2962 ATCAATGCTGTATGTATGTAACGTAATACACAGGCGCTTAACCTACCGTAAGAGGTTC 3021

OY 366 SerLySValPheAlaLySgIlyThrGlnMetIle-----AlaLySAsnProLySLeuSerLyVal 386
|          |||          |||          |||          |||
Db 3022 AGCATTAACGCAACAGCGCACCTGTGTATTAACGCAAAAGAGCGTGAAGTAAATGTG 3081

OY 377 -----AlaLySAsnProLySLeuSerLyVal 386
|          |||          |||          |||          |||
Db 3082 GAGGCATCAGTAAACCATACAGTACAGTGAATGCAACCAACGCAAAATGCTCCGCGAGCGTA 3141

OY 387 lIleSer-----SerLeuThrSerLySTrpValThrVal---GlyVal 399
|          |||          |||          |||          |||
Db 3142 ATCGGCAACCTCAAGCAGAGTGAACATCACTGGGGATTTAATACAAATAAATGAATTA 3201

```

Oy	295	-----AlaAlaAla	297
Db	2916	TACGGTTACGAGGGCGACGCCCTTGTGGCCCAAGCTGATCCGCACAGACGCGCCGCGAC	2975
Oy	298	GlycylAlaAlaGlyAlaAlaAlaAla---ThrThrValAlaThrGlnIleThrValGln	316
Db	2976	GGCGCGCTGCCCGCGCTTCGCCGCTCGACGACGAGTGAGGGCGCGGTGTCGATCGCC	3035
Oy	317	AlaValValGlnAlaLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAla	336
Db	3036	GCGCGCGCCAGCGTGA---GGGCGCCGGTGGCGGTACTCTTCTTGACGCGG	3092
Oy	337	AlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAla	356
Db	3093	GCCTCACAACGCATCTCTGGCGCGCTGCAGCACGCCCATCATGAAACACTGCCACGAT	3152
Oy	357	IleAlaLysAlaIleSerLys-----GlyIleSerLysValPheAlaLysGlyThrGln	374
Db	3153	TACGCTCGAAGATTGCACAGCTGGCGCGCCCAAGCGTACTGCAAAAACCTGAG	3212
Oy	375	-----MetIleAlaLysAsnProLysLeuSerLysValIleSerSerLeu	390
Db	3213	GCGCGTCACGAACAACCTGGC---AATTGGACAGGCTTAGGAAAAATCTGGCAGCTG	3269
Oy	391	ThrSerLysTrpValThrValGlyValValAlaAlaAlaAlaAlaAlaLeuGlyLys	410
Db	3270	CAGGCGCGTGGAGAACCAGCACGCTG---ATC	3299
Oy	411	GlyIleMetGlnMetGlnLeuSerGlyMetGlnGlnAsnValAlaGlnPheGlnLysLu	430
Db	3300	GGGGTGCAGACGACAGAGATCTCAAAGTGGCGGCTCGAAGTGGCGCCATTACGGCAC	3359
Oy	431	ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln	450
Db	3360	GCGGACAACCTGAATTCGTCGACGCTTCGTGAGACCGCTTCGACAGGCGAGGGGTG	3419
Oy	451	AlaSerLys-----IlaAlaSerLysGln	458
Db	3420	GCGGCGCAGCGGTGCTCTCGACGTCGCCCGCGCGCGCATGATATGCGACGCGCAG	3479
Oy	459	ThrGlyLysSerAsnGluMetThr-----	466
Db	3480	--GCGAGCGGCGCGGCTGACGTTCATCAGCAGCGCGCTGGCGCGCAGAGAACGAG	3536
Oy	467	GlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyAlaAlaIleSerGly	485
Db	3537	CGCGCGGCGACGAAACGGCGAAGACGGAATTCACACATTCGTGCAGATCGTGGCG	3593
RESULT 44			
US-09-702-251-3			
Sequence 3, Application US/09702251			
Patent No. 6372492			
GENERAL INFORMATION:			
APPLICANT: C. Frank Bennett			
APPLICANT: Lex M. Cowsett			
TITLE OF INVENTION: ANTISENSE MODULATION OF TALIN EXPRESSION			
FILE REFERENCE: RTS-0199			
CURRENT APPLICATION NUMBER: US/09/702,251			
CURRENT FILING DATE: 2000-10-30			
NUMBER OF SEQ ID NOS: 89			
SEQ ID NO 3			
LENGTH: 8078			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (127)..(7752)			
US-09-702-251-3			

[illegible]

Best Local Similarity:	23.74%	Mismatches:	198
Query Match:	6.23%	Indels:	81
DB:	4	Gaps:	16

US-09-889-314-2 (1-496) x US-09-702-251-3 (1-8078)

Oy	59	AlaThrIleAlaGlyAlaSerGlyLysAspLysPheSerSerThrThrLysThrglnThr	78
Db	1944	GCCTCTGGTGGAGGAGCAAAAGGCCCTTCGGGAGCACTGTGCAGACACTGTCCGCAATGC	2003
Oy	79	AlaProGlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAla	98
Db	2004	CCAACCAACCAGGAGTGTGAGCCCGCGACAAACCTGCTGCAGACAGCTGGGAACGTGG---	2060
Oy	99	AspThrGlyValSerGlyValAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla	118
Db	2061	--CCAGGCCAGTGGGAGCTGTTGGCAACAATT--GGGGAAGATGATACAGACCCCCAC	2115
Oy	119	MetGlnThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSer-----	136
Db	2116	TTCCAGGATGGCCTATATGACAGTGCACAAAGTGTGGCAAGTGTGGACGTGGACGTCCCTGTC	2175
Oy	137	LeuGlnSerLeuSerAlaAlaGlnMetLysGlu-----ValGlnAlaValAlaVal	153
Db	2176	CTCAAGGCGCAGAGTGTGGTGGCCCGAGCGAGACAGACGTGGGACCTTCACACCAACTGATT	2235
Oy	154	AlaAlaLeu-----SerGlyLysSerSerGlySerAlaLysLeuGlu	167
Db	2236	GCTGGAGCAACACAGTGTGCCCTATCCACTTCCCAACATGAGCCTGTACTAAGGTGGTGG	2295
Oy	168	ThrProGlnLeuProLysProGlyProGlyValThrProArgSerGluIleGluIleGlyLeu	187
Db	2236	GCACCTACAAATCAGCTCACT-----GTCTGCCAAGACCAACTGTGTGAGGCTGTGACGA	2349
Oy	188	AlaLeuAlaLysAlaAlaIleGlnThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsn---	206
Db	2350	CTGGTAGCCCAAGCCGTGGAGGGCTGTGTGTGTCTCCCAAGCGACGTACAGAGGATGGG	2409
Oy	207	-----TyrAlaSerThrGlnAla-----	212
Db	2410	CAACTGTGGCGAGGGGTGAGAGCAGCAGCCACAGCTGTACCCAGGCCCCTAATATGACTG	2469
Oy	213	-----GlnAla	214
Db	2470	CTGCAGCAGATGTCAAAAGCCATCCACAGAGGGCTGGGCTGCTGGCCGTTATGACCAAGCT	2529
Oy	215	AspGlnThrAsnLysLeuGlnGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlu	234
Db	2530	ACTGACACCAATCCTAACCGTCACTGAGAAACATCTTACCTCCATCGATGCTCTGGGAG	2589
Oy	235	GluTyrGlnGlnMetLysAlaAlaIleGlnLysSerLysAspLeu-----	249
Db	2590	ATGTGGGAGAGCCCGGACATCTGTGGCCCAAGCCACATCTGACCTGCTCAATGCCATCAG	2649
Oy	250	-----GlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAla	266
Db	2650	GCTGATGCTGAGAGGGGAAGTATCTGGAGAACTCCCGCAACTC---TTAAGTCTCTGC	2706
Oy	267	IleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu	286
Db	2707	AAGATCTCAAGCTGATGATGCCACACCAAGATGTATGAGAGGTGCCAAGGAGCACTGCCAC	2766
Oy	287	AlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaGlyAlaAla	306
Db	2767	CCTGCACAGTGA--GGAGCAGCAGCAGCAGCGGCTGGGAGGACGTGAGGGGCTGTGCGATGCC	2825
Oy	306	aThrThrValAlaThrGlnIleThr-----ValGlnAlaValAlaGlnAla	321
Db	2826	CACCAATGACGTGGCGCAATGATGCCATCAAGAAAGCTGTGGCAGCGCTGCAGCATATGC	2885
Oy	321	aValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAla	341
Db	2886	ACCCAAGAGGTGTGACGCTCTGCCACA---CAGACATCGCTGTGAGCTACAGCAGCAGCG	2942

QY	294	-----	294	-----	294
Db	2854	CTGGTCAGCAATGCACTATGCGATCAGCTGGACAAAGCTGGCCAGGAATCGAGCGCA	2913		
QY	295	-----	295	-----	295
Db	2914	TACGTTACAGAGCGCGACCGCTTGGTGGCCAGCGTATCGCAGAAAGCGCGCCGAG	2973		
QY	298	GIYGIYAlaAlaGIYAlaAlaAlaA---ThrThrValAlaThrGlnIleThrValGln	316		
Db	2974	GGCGCGCGTGGCGCGGTCTCCCGCTCTCTGACAGAGTGGGGCGGGGTCTGCATGCC	3033		
QY	317	AlaValAlaGlnAlaValAlaGlnAlaValAlaIleThrAlaValArgGlnAlaIleThrAla	336		
Db	3034	GGCGCGCCAGCGGTGTA---GGGGCCCCGGTGGCGGTGATCTCTTGGTACCGGG	3090		
QY	337	AlaIleAlaAlaAlaValAlaValAlaValSerGlyIleValAlaPheIleThrIleValAla	356		
Db	3091	GCCTCAGACGCGATCCTCGTGGCGCGCTGACAGCCCATCATGAAAGCTGGCCACAGT	3150		
QY	357	IleAlaAlaAlaIleSerLys-----GlyIleSerLysValPheAlaLysGlyThrGln	374		
Db	3151	TACGCTGCAAGATTCAGACGAGCGTGGGGCGGCCAGCGCTACTTCAGAAAACTCGCG	3210		
QY	375	-----MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeu	390		
Db	3211	GCGGTCAGCAACACACGAGCC---AAATTCGACGCGCTTACGAAATGCTGGCGCACGTG	3267		
QY	391	ThrSerLysTrpValThrValGlyValAlaValAlaAlaProAlaLeuGlyLys	410		
Db	3268	CAGCGCGGTTGAGACGCGACGAGCGTG-----ATC	3297		
QY	411	GlyIleMetGlnMetGlnLeuSerGlnIleMetGlnIleAsnValAlaGlnPheGlnLysGln	430		
Db	3298	GGGGTGGCAGACAGACAGAAATCTCCAAAGTCGGCGCTCGAAGTGGCGGCATTACCGCAAC	3357		
QY	431	ValGlyLysLeuGlnAlaAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGln	450		
Db	3358	GGCGACAACCTGAATCGCTGAGCTGTCTGACGCGCTTCGTCAGGGCGGAGGGGCTG	3417		
QY	451	AlaSerLys-----IleAlaSerLysGln	458		
Db	3418	GGCGGCGACGCGCGGTGCTCTCGACGTCCGCCCGCGCGGCATCATATTCGACGCGCAAG	3477		
QY	459	ThrGlyLysAsnGlnIleMetThr-----	466		
Db	3478	---GGCGACGGCGCGCGCTGACGTTCAACACGCCGTCGGCGCGCCAGAGAGAGCAG	3534		
QY	467	GlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyAlaAlaAlaIleSerGly	485		
Db	3535	CGCGCGCGCGCAAAAGGGCAAGAGCGAATTCACCACTTCGTCGAGATCGTGGGC	3591		
RESULT 43					
5183745-5					
; Patent No. 5183745					
; APPLICANT: DANCHIN, ANTOINE;GLASER, PHILIPPE;KRIN, EVELYN;					
; BARZO, OCTAVIEN;LADANT, DANIEL;ULMAN, AGNES					
; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR					
; BIOLOGICAL USES					
; NUMBER OF SEQUENCES: 13					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/426,541					
; FILING DATE: 25-OCT-1989					
; SEQ ID NO:5:					
; LENGTH: 6443					
5183745-5					
Alignment Scores:					
pred. No.:		0.000204	Length:	6443	
Score:		147.50	Matches:	114	
Percent Similarity:		32.47%	Conservative:	74	
Best Local Similarity:		19.6%	Mismatches:	210	

Query Match:	6.23%	Indels:	181
DB:	6	Gaps:	22
US-09-889-314-2 (1-496) x 5183745-5 (1-6443)			
QY 12 ProasphanGlnLysAsnIleMetSerGlnValThrSerThrProGln-----			28
DB 2085 CCGACGGACGCTTCGAAGTTCCTCCGCCGAGATCACTGAAACGGGTCCGCCCTCACCCGGA			2144
QY 29 -----GlyValProGlnIAspLysLeuSerGlnValThr			41
DB 2145 TTGGCGCGCGCTCGCTGGCGCGCTGGAGAACGCCAGAT-----TCCGGCTATGACAC			2198
QY 42 LysGlnIleGlnIleThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIle			61
DB 2199 CTTGATGGGGTGGATCCGATCCGATCTTCGTTGGCGAGGTG---TCCGACATGGCCGCC			2255
QY 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln			81
DB 2256 CTGGAGCGCGCGAAGCTGGAAATGACC-----CGG			2285
QY 82 GlnGlyValAlaAlaGlyLysGluSerSerGlnLysAlaGlyAlaAspThrGly			101
DB 2286 CAAGTCTTCGACGCCGGGGCGCGCAGACAT-----GCCAGCCGGCG			2330
QY 102 ValSerGlyAlaAlaAla-----Thr			108
DB 2331 GTGAGCGGTCTCGCGCGCACTGGGGCGACGCGGCGCTGCAGGGCGCCACGGCTGGCG			2390
QY 109 ThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnAlaSerLys			128
DB 2391 GCGGCGACGGCGCTGCTTCATGCGCATTTGCCCTGATGACCAATTCGGCGCGCGGTTC			2450
QY 129 SerMetGluSerThrLeuGlnIleGlnSerLeuGlnSerLeuSerAlaAla-----GluMetLys			146
DB 2451 ACC-----AACACGGCGCGAGAGCGGCGCTCTTCTGCGCGCGCTGTGGCGCTGGCG			2504
QY 147 GluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeu			166
DB 2505 GAGGCGACGACGCGCGCTGGCGCAACCGTAGCGGTTTTTTCGCGCGCTTCGCGCTGG			1564
QY 167 GluThrProGluLeuProLysProGlyValThrProLysSerGluValIleGluIleGly			186
DB 2565 GCCGCGC-----GGT			2573
QY 187 LeuAlaLeuAlaLysAlaIleGlnIleThrLeuGlyGluAlaThrLysSerAlaLeuSerAsn			206
DB 2574 TTGCGCGTGCCTGGCGCGCGCATGCGCTGGAGAGCGCATCCGCCCGCGCTGGC---			2630
QY 207 TyrAlaSerThrGlnAlaGlnAlaAspIleThrAsnLysLeuGlyLeuGlnAla			226
DB 2631 -----GCCGGG			2636
QY 227 IleValIleAspLysGluArgGlnGluIleThrGlnGluMetLysAlaAlaGlnLysSer			246
DB 2637 ATCTGCTGTGACGATGACCGCGCGCGCAGAGAGAGCGCGCGCGCGCGCGCATGCC			2696
QY 247 LysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAla			266
DB 2697 CTCACAGTTGACAGT-----GGAACGGTCGAGCGTCTTCATCGCGC			2741
QY 267 IleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu			286
DB 2742 TTGGCGCTGGCGCGC-----GCCGCGCGCTGACACAGCGGCTTGACAGTGGCGCGGGCG			2795
QY 287 AlaAlaGlyAlaAlaValAlaGlyAla-----			294
DB 2796 TCGCGCGGGCGGCTCGCGCGCATTTGGCCCGCGCGCTCACTCCATGAGCATGACGCG			2855
QY 294 -----			294
DB 2856 CTGTGTCAGCAATCGCATCTATGGGATCACTGGACAGCTGGCGCAGGATCGAGCGCA			2915

```

Db 2473 CAGCGCGGTGACAGCCAGCAGCGTG-----ATC 2502
Oy 411 G1y1IeMeTg1nMeSg1nLeuSg1nMeTg1nG1nAsnVala1aG1nPhg1nLys1u 430
Db 2503 GGGGTGACAGCAGACAGATCTCCAGTCGGCGCTCGACATTCGACCGGCAAC 2562
Oy 431 ValG1yLysLeuG1nAla1aAla1aAspMeT1LeSerMeTPhg1nPhg1n 450
Db 2563 GCGGACACCTGGAATCCGTGACGTGTTCGTGGACCGCTTCGTCCAGGGAGCGGTG 2622
Oy 451 AlaSerLys-----1La1aSerLys1n 458
Db 2623 GCCGCGCAGCGGTGCTCTCGACGTCCGCCGCGGCATCGATATCGCAGCGCGCAAG 2682
Oy 459 ThrG1yLysAsn1uMeThr----- 466
Db 2683 --GGCGAGCGCGCGCTGCTGACGTCATCACGCCGTCGCCGCCAGAGAGAGAG 2739
Oy 467 G1nLysAlaThrLysLeuG1yAlaG1n1LeuLysAlaTyrAla1a1eSerGly 485
Db 2740 CGCCGCGCAGAAAGCGGCGCAGAGCAATTCACACATTCGTGACATCTGTGGC 2796

```

## RESULT 42

```

US-08-669-785-1
: Sequence 1, Application US/08669785
: Patent No. 6309648

```

```

: GENERAL INFORMATION:
: APPLICANT: Betsou, Potnl
: APPLICANT: Sebo, Peter

```

```

: APPLICANT: Guiso, Nicole
: TITLE OF INVENTION: Protective Epitopes of Adenyl

```

```

: TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
: TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.

```

```

: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESS: Finegan, Henderson, Farabow, Garrett & Dunner
: STREET: 1300 I Street, N.W., Suite 700

```

```

: CITY: Washington
: STATE: D.C.

```

```

: COUNTRY: USA
: ZIP: 20005-3315

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:

```

```

: APPLICATION NUMBER: US/08/669,785
: FILING DATE: 27-JUN-1996

```

```

: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:

```

```

: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146

```

```

: REFERENCE/DOCKET NUMBER: 02356, 0072-00000
: TELECOMMUNICATION INFORMATION:

```

```

: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400

```

```

: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 6441 base pairs
: TYPE: nucleic acid

```

```

: STRANDEDNESS: double
: TOPOLOGY: linear

```

```

: MOLECULE TYPE: DNA (genomic)
: FEATURE:

```

```

: NAME/KEY: CDS
: LOCATION: 979..6096

```

```

: OTHER INFORMATION: /note="AMINO ACID SEQUENCE
: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE

```

```

: OTHER INFORMATION: CODING FOR THE B. Pertussis AC-Hly"
: US-08-669-785-1

```

## Alignment Scores:

```

Pred. No.: 0.000204 Length: 6441
Score: 147.50 Matches: 114
Percent Similarity: 32.47% Conservative: 74
Best Local Similarity: 19.69% Mismatches: 210
Query Match: 6.23% Indels: 181
DB: 4 Gaps: 22

```

```

US-09-889-314-2 (1-496) x US-08-669-785-1 (1-6441)

```

```

Oy 12 ProAspAsnG1nLysAsn1LeMeSerG1nValLeuThrSerThrProG1n----- 28
Db 2083 CCGAGCGGACGTTGCAAGTTCTCCCGGATGTACTGGAAAGCGTCCGCGTCAACCGGA 2142
Oy 29 -----G1yValProG1nG1nAspLysLeuSerG1yAsnG1nThr 41
Db 2143 TTGCGCGCGCGCTGCTGCGCGGAGTGAAGCCAGAT-----TCCGACATGCGCGCC 2196
Oy 42 LysG1n1LeG1nInThrArgG1nG1yLysAsnThrG1uMeTGlusSerAspAlaThr1le 61
Db 2197 CTTGATGGGCGTGGATCCGATCGTTCTGTTGGCGAGATG---TCCGACATGCGCGCC 2253
Oy 62 AlaG1yAlaSerG1yLysAspLysThrSerSerThrThrLysThrG1uThrAlaProG1n 81
Db 2254 GTGAAGCGCGCGCACTGGAATGAC-----CGG 2283
Oy 82 G1nG1yValAlaAlaG1yLysG1uSerSerG1uSerG1nLysAlaG1yAlaAspThrGly 101
Db 2284 CAGTCTTGACAGCGCGCGCGCGCGAGACAT-----GCCAGCGCGGCG 2328
Oy 102 ValSerG1yAlaAlaAla-----Thr 108
Db 2329 GTGACCGGCGCTCGCGCACTGGGCGAGCGCGCTGCAGGCGCCAGCGCGTGGCG 2388
Oy 109 ThrAlaSerAsnThrAlaThrLys1LeAlaMeTg1nThrSer1LeG1uLysLeuSerLys 128
Db 2389 GCGGCGACGCGCTGCTCATGCCATGCTGATGACGCAATTCGGCGCGCGCGCTTC 2448
Oy 129 SerMeTGlusThrLeuG1uSerLeuG1nSerLeuSerAlaAla-----G1nMeTlys 146
Db 2449 ACC-----AACACGCGCGAGAGAGCGGCGCTGTTGCGCGCGCGTGTGGCTTGGCC 2502
Oy 147 G1uValG1uAlaValAlaValAlaAlaLeuSerG1yLysSerG1ySerAlaLysLeu 166
Db 2503 GAGCGCAGCAGCGCGTGGCGCAACCGGATGCTTTTCGCGGCGCTTCGCGCTGG 2562
Oy 167 G1uThrProG1uLeuProLysProG1yValThrProArgSerG1uVal1LeG1u1eGly 186
Db 2563 GCCGCG-----GCT 2571
Oy 187 LeuAlaLeuAlaLysAla1LeG1nThrLeuG1yG1uAlaThrLysSerAlaLeuSerAsn 206
Db 2572 TTGCGCGTGGTGGCGCGCGCATGGCGTGGAGCGGATGCGCGCGCGCTTGGC--- 2628
Oy 207 TyrAlaSerThrG1nAlaG1nAlaAspG1nThrAsnLysLeuG1yLeuG1uLysG1nAla 226
Db 2629 -----GCCGCG 2634
Oy 227 1LeLys1LeAspLysG1uArgG1uTyrG1nG1uMeTlysAlaAlaG1uG1nLysSer 246
Db 2635 ATGTGCTTGAACGAGTACGCGCGCGCGCGAGCAAGAACCGCGCGCGCGCGCGATCGCG 2694
Oy 247 LysAspLeuG1uG1yThrMetAspThrValAsnThrValMet1LeAlaVala1a 266
Db 2695 CTGCAATTGACAGT-----GGAACGCTGAGCTGCTTCTTCATCGCGG 2739
Oy 267 1LeThrVal1LeSer1LeValAlaAla1LePheThrCysG1yAlaG1yLeuAlaG1yLeu 286
Db 2740 TTGGCGCTGGCGCG-----GCCGCGCGCTGACACAGCGCTTCAGCGTGGCGCGCGG 2793
Oy 287 AlaAlaG1yAlaAlaValaG1yAla----- 294
Db 2794 TCGCGCGCGCGCTCGCGCGATTCGCGCGCGCGCTCATGTCATGAGATCTACGGC 2853

```





```

Db 2824 -----CTGAGAGAGCTCTAGC 2840
Oy rleuSerAlaIaIaGlnMetLysGluValGluValAlaValAlaIaIaIeSerGly 159
Db 2841 TGTGGAGCGGAGCGGCTTCTCTATGAGGTGTTCCAGCTGCGGCGGTGAGCGGC 2900
Oy sSerSerGlySerAlaLysLeuGluThrProGlu-----LeuProLysProGlu 175
Db 2901 TCCGGTGGGGGCTCA-----AGTACTCCGAGCGGCGGCTCTAGAGCGGCGGC 2954
Oy 175 yValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnTh 195
Db 2955 GCGTACTCCGGAGCGGCTGTG-----GCTCTAGCTGGCGGCTGGCTTAC 2999
Oy rleuGlyGluAlaThrLysSerAlaLeuSerAsnTyraSerThrGlnAlaGlnAlaAs 215
Db 3000 TCTGGGGCGGCGGCGCTCCAGCT---GCGAGGTGCTCTCCGAGGCGGCGCG- 3054
Oy 215 pGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGlu 235
Db 3054 ----- 3054
Oy 235 uTyrGlnLysMetLysAlaIaGlnLysSerLysAspLeuGluGlyThrMetAspTh 255
Db 3055 -----GCTCCAGCTGCGGCGGCGGCTCTCCGAGGCGGCTCTCCAGC 3098
Oy rValAsnThrValMetLysAlaIaValSerValAlaIleThrValIleSerIleValAla 275
Db 3099 TGTGGAGCGGCGGCGCTCTGCTGGGCGGCTCAAGTACTCCGAGGTGCGGCGGC 3158
Oy 275 AliePheThrGlyGlyAlaGlyLeuAlaGlyLeuAlaIaGlyAlaIa----- 291
Db 3159 GCGCT-AGCTCGGCGGCGGCTCTCCGAGGCGGCGGCGGCTCTCCAGCTGCGGAG 3217
Oy 292 -----ValGlyAlaAlaAlaAlaGlyGlyAlaAlaGly-A 303
Db 3218 CGGATCAGAGCGGCGGCTCTCTGAGGCGGAGCTCTGAGGCGGAGCTCGGCTG 3277
Oy 303 lAlaAlaIaIaThrThrValAlaIaThrGlnIleThrValGlnAlaValAlaVal 323
Db 3278 CTGCGGCGGCTTCCGAGGCGGCAAGTACTCTGCTGGCGGCGGCTCCAGCTGCGG 3337
Oy 323 yGlnAlaValIleThrAlaIaArgGlnAlaIleThrAlaIaIleLysAlaVal 343
Db 3338 AGCGGCTATTCGCGGCGGCTGGAAGCAGCGGCGGCTCTCTGCGGCGGCTACTC 3397
Oy 343 ySerGlyIleLysAlaIaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 363
Db 3398 AGGTGCGGCTGATCA-----GCTGCGGCGGCGGCGG 3430
Oy 363 yGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 383
Db 3431 CGGCTATTCGCTGCGGCGGAGCAGCTCGGAGGTGTT-----CCTCCG 3478
Oy 383 eSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403
Db 3479 TGGCGGCGGCGGCGATGCTCCCAACAGTATCAGTGCCAGACTACGAGGCGTCTAG 3538
Oy 403 aAlaAlaProAla 407
Db 3539 CGGTGCTCCAGCT 3552

```

RESULT 40  
 US-08-961-527-377  
 ; Sequence 377, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue

```

; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 377:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-377

```

```

Alignment Scores:
Pred. No.: 5,44e-06 Length: 690
Score: 147.50 Matches: 62
Percent Similarity: 37.23% Conservative: 43
Best Local Similarity: 21.99% Mismatches: 112
Query Match: 6.23% Indels: 65
DB: 4 Gaps: 5

```

US-09-889-314-2 (1-496) x US-08-961-527-377 (1-690)

```

Oy 56 GluSerAspAlaThrIleAlaGlyLysSerGlyLysAspLysThrSerThrThrLys 75
Db 33 GAATCGGATCAACAGAGTCTTCCGCTTCCAGCATCAGACAGAGCGGCTTCCAGCAAC 92
Oy 76 ThrGluThrAlaProGlnGlnGlyValAlaIaGlyLysGluSerSerGluSerGlnLys 95
Db 93 ACATCAGCTTCTGAATCTGCATCAACCACTGCGCTTCCAGCGTCAACCACTGCGTGC 152
Oy 96 AlaGlyAlaAspThrGlyLysSerGlyAlaAlaIaIaThrThrAlaSerAsnThrAlaThr 115
Db 153 GCTTACGCTCGACACAGTCTTCCGCTTCCAGCATCAGACAGTCTTCCAGCAAC 212
Oy 116 LysIleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGlu 135
Db 213 -----GCAAGTACTCAGCGTCACTTCCGCTCA 242
Oy 136 SerLeuGlnSerLeuSerAlaIaGlnMetLysGluValGluAlaValAlaValAla 155
Db 243 ACCAGTGCCTCCGCTTCCAGCAACAGTGCCTCAGCTCAGCAAGTATTCACAGCTCT 302
Oy 156 LeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGly 175
Db 303 GAATCGCATCAACAGTCTGCGGCTCA----- 332
Oy 176 ValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195
Db 333 -----GCAAGCGCAGTACTCA 350
Oy 196 LeuGlyGluAlaThrLysSerAlaLeuSerAsnTyraSerThrGlnAlaIaIaAsp 215
Db 351 GCGTACGCTTCCGCTCAACCACTGCGTCTTCCAGCAAGCAAGTCTGCTCAGCTCA 410

```

```

OY 99 pthrglyValserGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMe 119
Db 2787 TATAGCGGTGGCGCTCTGGCTCCGAGCGGCTCAT----- 2823
OY 119 tGlnThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSe 139
Db 2824 -----CTGGAGGAGGCTCTAGC 2840
OY 139 rLeuSerAlaAlaGluMetLysGluAlaAlaValAlaValAlaAlaLeuSerGlyLy 159
Db 2841 TGTGGAGCGGAGCGGTGGTCTCTATGAGAGTGTTCCAGCTGCGCGGTGGAGCGCGC 2900
OY 159 sSerSerGlySerAlaLysLeuGlnThrProGlu-----LeuProLysProG 175
Db 2901 TCCCGTGGGGGCGTCA-----AGTACTCCGAGAGCGCGGTGGCTACTGCGCGCGC 2954
OY 175 yValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnTh 195
Db 2955 GGTACTCTCCGAGCGCGGTGTG-----GCTCTACTGCGCGGTGGCTAGC 2999
OY 195 rLeuGlyGluAlaThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAs 215
Db 3000 TCTGGGGGCGCGCGCGCTCCAGCT---GCGAGAGTGCGTACTCCGAGAGCGCGCGG-- 3054
OY 215 pGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluG 235
Db 3054 ----- 3054
OY 235 uTyrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGlnGlyThrMetAspTh 255
Db 3055 -----GCTCCAGCTGCGGCGCGCGCTACTCTCCGAGGTGGCTCCAGC 3098
OY 255 rValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAl 275
Db 3099 TGTGAGCGCGGTGGCGCTCTGTGGGGCGCGTCAAGTACTCCGAGAGTGTGGCGCGCGC 3158
OY 275 aIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAla----- 291
Db 3159 GAGCTCT-AGCTGCGCGCGCGGCTCTCCGCGGCGCGCGCTCCAGCTGCGGAGG 3217
OY 292 -----ValGlyAlaAlaAlaAlaGlyGlyAlaAlaGly-A 303
Db 3218 CGGATCAGAGCGCGCGCTCTACTGCGAGAGCTCTCTGAGAGGCGAGCTCCGCTGCG 3277
OY 303 lAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValL 323
Db 3278 CTGGCGCGCGGTCCGAGCGCGCAAGTACTGTGGGGCGCGGTCCAGCTGCGG 3337
OY 323 ySglnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValL 343
Db 3338 AGCGGCTATTCGCGCGCGGTGAGACAGCGCGCTCTAGCTGTGGCGCGCGCTACTC 3397
OY 343 ySserGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerL 363
Db 3398 AGGTGGCGGTGATCCA-----GCTGCGCGCGCGCGCG 3430
OY 363 ySgylIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysL 383
Db 3431 CGGCTATTCGCGCGCGGTGAGACAGCTGCGAGAGTGTGTT-----CTCCCGG 3478
OY 383 eUerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValValVal 403
Db 3479 TGGCGCGCGCGCGGTCTCCCAACAGTATCACTGCCAGAGCTACGAGGCGGTCTTAG 3538
OY 403 aAlaAlaProAla 407
Db 3539 CGGTGGCTCCAGCT 3552

```

RESULT 39  
 PCT-US93-03993-1  
 ; Sequence 1, Application PC/TUS9303993  
 ; GENERAL INFORMATION:

```

APPLICANT: Roop, Dennis R.
APPLICANT: Rochnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL
TITLE OF INVENTION: VECTOR SYSTEMS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Fulbright & Jaworski
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03993
FILING DATE: 19930428
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5405
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5225
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US93-03993-1

Alignment Scores:
Pred. No.: 0.000185 Length: 6530
Score: 148.00 Matches: 92
Percent Similarity: 33.74% Conservative: 45
Best Local Similarity: 22.66% Mismatches: 165
Query Match: 6.25% Indels: 106
Gaps: 10
DB:

US-09-889-314-2 (1-496) x PCT-US93-03993-1 (1-6530)
OY 37 serglyAsnGluThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGlu 56
Db 2548 AATGAGTCCAGACACCTAGTGTGATTAAGCTAGAGTCAAGAAATGTGTGGCATGGAT 2607
OY 57 SerAspAlaThrIleAla----- 62
Db 2608 GCATCTGCACCTTACAGAGCTCTGCTGCTGTGGTCTAATGTGCTCTTCTGCTC 2667
OY 63 -----GlyAlaSerGlyLysAspLysThrSerThrThrLysThrGlnThrAla 79
Db 2668 TTCTTCCAGCGTTCCTCTCTCTTA-AAACAAGATCTCACCAAAAGACAGCCACT 2726
OY 80 ProGlnGlnGlyValAlaAlaGlyLysGluSerSer-GluSerGlnLysAlaGlyAlaAs 99
Db 2727 CCTTGCCTCTCTGTGGGTGTGGAAAGACCTCTGTGAGAGGAGGAGCGCGCGCTAT 2786
OY 99 pthrglyValserGlyAlaAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMe 119
Db 2787 TATAGCGGTGGCGCTCTGGCTCCGAGCGGCTCAT----- 2823
OY 119 tGlnThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSe 139

```

```

Db 2727 CCCTGCTCTGCTGTTGGAAAGACCTCTGTCGAGAGAGCGCGCGCTAT 2786
Oy 99 pthrclyvalserglyalaalaThThralaSerasnThralaThrllyleAlaMe 119
Db 2787 TATACGGGTGGGCTCGGCGGAGCGGCTCAT----- 2823
Oy 119 tclnThSerllegluAlaSerlySerMetcluserThrlengluSerleugluse 139
Db 2824 -----CTGGAGGAGGCTCTAC 2840
Oy 139 rleuserAlaalaIngluMetlySgluValgluAlaValaValaAlaalaLeuserglyly 159
Db 2841 TCTGAGCGCGAGCGGCTCTCTATGAGTCGTCACGTCGCGCGCTGAGAGCGGC 2900
Oy 159 sserSerglySerAlaLySleugluThProglu-----leuProLyProgl 175
Db 2901 TCCGGTGGGGGCTCA-----AGTACTCCGAGAGCGCGGCTGCTCTACGCGCGC 2954
Oy 175 yvalThrProArSergluVallllegluilegluAlaalaAlaAlaAlaAlaAla 195
Db 2955 GGCTACTCCGGAGCGGCTGCTG-----GCTCTAGCTGCGCGCGCTGCTAC 2999
Oy 195 rleuglygluAlaThrLySserAlaLeuserAsnTyraLaserThrglnAlaAlaAs 215
Db 3000 TCTGGGGCGCGCGGCTCCAGCT---GCGAGGTGCTGCTCCGAGAGCGCGCGC-- 3054
Oy 215 polnThrlnlySleuglyleugluSgluAlaAlaAlaAlaAlaAlaAlaAla 235
Db 3054 ----- 3054
Oy 235 uTyrglnluMetlySalaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 255
Db 3055 -----GCTCCAGCTCGCGCGCGCGCGCGCTACTCCGGGGGCTGCCAGC 3098
Oy 255 rValasnThrValMetlleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 275
Db 3099 TGTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3158
Oy 275 alleTherThrySgllyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 291
Db 3159 GCTCT-AGCTCGCGCGGCTCTCGCGCGCGCGCGCGCGCGCTCCAGCTCGCGAGG 3217
Oy 292 -----ValglYAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 303
Db 3218 CGGATCAGAGAGCGCGCTCTACTGCGGAGGCTCTTGAGGCGGAGCTCCGCTG 3277
Oy 303 laAlaAlaAlaThrThrValAlaThrglnleThrValglAlaAlaAlaAlaAla 323
Db 3278 CTGGCGCGCGGCTCGGAGCGCAAGTCTGTGCGCGCGCGCTCCAGCTGCGG 3337
Oy 323 ySglAlaValleThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 343
Db 3338 AGCGCGCTATTCGCGCGCGGAGCAAGCGCGGCTCTACTGCTGCGCGGCTACTC 3397
Oy 343 ySserglyleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 363
Db 3398 AGCTGCGCGTGATCA-----GCTGCGCGCGCGCGCGG 3430
Oy 363 ySglYleSerlyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 383
Db 3431 CGGCTATTCGCTGCGCGGCGAGAGCTGCGAGGTGCT-----CTCCCG 3478
Oy 383 euSerlyValleSerSerleuThrSerlyTrpValThrValglYAlaAlaVal 403
Db 3479 TGGCGCGCGCGGAGATCTCCCAACAGTATCAGTCCAGAGCTACGAGGCGGTTCTAG 3538
Oy 403 aAlaAlaAlaProla 407
Db 3539 CGGTGCTCCAGCT 3552

```

```

; Patent No. 6143727
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: LYON & LYON
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,240
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,930
; FILING DATE:
; APPLICATION NUMBER: 07/876,286
; FILING DATE: April 30, 1992
; APPLICATION NUMBER: No. 6143727 yet assigned (204/132)
; FILING DATE: October 29, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6530 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-458-240-1
;
; Alignment Scores:
; Pred. No.: 0.000185 Length: 6530
; Score: 148.00 Matches: 92
; Percent Similarity: 33.74% Conservative: 45
; Best Local Similarity: 22.66% Mismatches: 165
; Query Match: 6.25% Indels: 106
; DB: 3 Gaps: 10
;
; US-09-889-314-2 (1-496) x US-08-458-240-1 (1-6530)
;
; Oy 37 SerglyAsnGluThrLySglInlleGlnThrArgGlnGlyLySAsnThrglnuMetGlu 56
; Db 2548 AATGAGGTCACACACCTGATGCTGATTAAGCTAGAGTACAGAAAGTGTGCGATGAT 2607
; Oy 57 SerAPAlaThrlleAla----- 62
; Db 2608 GCATCTGCACCTTCACACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2667
; Oy 63 -----GlyAlaSerGlyLySAspLySThrSerSerThrThrLySThrglnThrala 79
; Db 2668 TTCTTCCAGAGTTCCCTCTCTCTTA-AACAAGATGTCACACAGAAAAAGACGCCACT 2726
; Oy 80 ProGlnGlnGlyValAlaAlaAlaGlyLySgluSerSer-GluSerGlnLySAlaGlyAlaAs 99
; Db 2727 CCCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2786

```



280 GlyAlaGlyLeuAlaGlyLeuAlaGlyAlaAlaValAlaGlyAlaAlaGlyGly 299  
1391 AGCATATGATGAGTGGC---TCACACAGTACAGAGCTTTCAGACTCACAAGTACATCG 1335  
300 AlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 319  
1334 ACATCAGACAGTACATGCTGATCACA---AGTGAAGAGGAGCTCGGACAGTGAAGT 1281  
320 GlnAlaValLysGlnAlaValLysAlaValLysAlaValLysAlaValLysAlaValLys 339  
1280 ACGTCATCAAGTACATGATGAGTACAGAGCTTTCAGATTCACAAGCCGCTCAACGTCA 1221  
340 AlaAlaValLysSerGlyLysAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 359  
1220 GAAAGTGGCTCAACATCAACAGAGTGTG-----AGTACTCGAATATGCTCAAGTACG 1170  
360 AlaLeuSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAla----- 377  
1169 TCATTAAAGTATGATGACAGATGAGTCTTTCAGATTCACAGATGATGCTGACATCAGAC 1110  
378 -----LysAsnPheProLysLeuSerLysValLysSer 389  
1109 AGTGGTCACATCAACAAGTACAGAGCTGACAGTCAAGTACATGTTAAGTGAAC 1050  
390 LeuThrSerLysTrpValThrValGlyVal-----ValValAla 404  
1049 TCGACAAAGTACAGAGCTTTCAGATGACAGTACATCAACATCAACAGATGAGTGGC 990  
405 AlaProAlaLeuGlyLysGlyLysMetGlnMetGlnLeuSerGlnMetGlnAlaVal 424  
989 TCAACACAGTACAAAGTACATTAAGTACGACAGTACAGATGCTTTCAGATCAACAG 930  
425 AlaGlnPheGlnLysGlyValGlyLysLeuGlnAlaAla---AlaAspMetLysSerMet 443  
939 ACCACATGACATCAAGTACAGTACAAAGTACAGAGGAGGAGGAGGAGGAGGAGTGA 870  
444 PheThrGlnPheThrGlnAlaLysSerLysIleAlaSerLysGlnThrGlyLysSer 462  
869 AGTACATCTTAAAGTACAGTACAAAGTACATTCAGACTCAACAGATCA 813

RESULT 36  
US-08-669-785-3  
: Sequence 3, Application US/08669785  
: Patent No. 6309648  
: GENERAL INFORMATION:  
: APPLICANT: Betsou, Fotini  
: APPLICANT: Sebo, Peter  
: TITLE OF INVENTION: Protective Epitopes Of Adenyl  
: TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To  
: TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner  
: STREET: 1300 I Street, N.W., Suite 700  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20005-3315  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/669,785  
: FILING DATE: 27-JUN-1996  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Meyers, Kenneth J.  
: REGISTRATION NUMBER: 25,146  
: REFERENCE/DOCKET NUMBER: 02356, 0072-00000

TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202) 408-4000  
: TELEFAX: (202) 408-4400  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 5118 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: DNA (genomic)  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 1..5115  
: OTHER INFORMATION: /note= "AMINO ACID SEQUENCE  
: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE  
: OTHER INFORMATION: CODING FOR THE B. Bronchiseptica AC-Hly"  
US-08-669-785-3

Alignment Scores:  
Pred. No.: 7,83e-05 Length: 5118  
Score: 150.00 Matches: 116  
Percent Similarity: 33.04% Conservative: 74  
Best Local Similarity: 20.17% Mismatches: 211  
Query Match: 6.33% Indels: 174  
DB: 4 Gaps: 22

US-09-889-314-2 (1-496) x US-08-669-785-3 (1-5118)

QY 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27  
DB 1117 TCGAAGTCCCTCCGCGGATGCTACTGGAACGGTCCGCGCTACACCGGATTCGGCGCGCG 1176  
QY 28 Gln-----GlyValProGlnLysAspLysLeuSerGlyAsnGlnThrLysGlnIleGln 45  
DB 1177 TCGCTGGCGCGAGTGAACGACGAGTAT-----TCCGCTATGACAGCCTTGATGGCGTG 1230  
QY 46 GlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAlaSer 65  
DB 1231 GGATCGCGATTCCTTCCTGTTGGCGCAGTGC---TCCGATATGGCGCGCTGGAACGGCGG 1287  
QY 66 GlyLysAspLysThrSerSerThrLysThrGlnThrAlaProGlnGlnValAla 85  
DB 1288 GAATCGGAATATGACC-----CGCAAGCTTTCAC 1317  
QY 86 AlaGlyLysGlnSerSerLysSerGlnLysAlaGlyAlaAspThrGlyAlaSerGlyAla 105  
DB 1318 GCCGGCGCGCGGACGAGCAT-----GCCGACCGGCGCTGAGCGGTGCG 1362  
QY 106 AlaAla-----ThrAlaSerAsn 112  
DB 1363 TCGGCGCACTGGGGCGCGGCGCTGCAAGGCGCCAGCGGCGGTGGCGCGCGACGCG 1422  
QY 113 ThrAlaThrLysIleAlaMetGlnThrSerIleGlnAlaSerLysSerMetGlnSer 132  
DB 1423 CTGGTTCAATGATGCTGCTGATGACGCAATTCGGCGGCGGTTCACAC-----AAC 1476  
QY 133 ThrLeuGlnSerLeuGlnSerLeuSerAlaAla-----GlnMetLysGlnValGlnAla 150  
DB 1477 ACGCGCGCAGGAAGCGCGCTGTTGCGCGCGCTGCTTGGCGGCGGACGACACAC 1536  
QY 151 ValValAlaAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlu 170  
DB 1537 GCCGTGCGCGAACCCTGAGCGGTTTTCGCCGGGTCTTCGCCGTGGCGCGCG----- 1590  
QY 171 LeuProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAla 190  
DB 1591 -----GGTTTCGGCGTGGCT 1605  
QY 191 LysAlaIleGlnThrLeuGlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThr 210  
DB 1606 GCGCGCGCGATGCGCTGGGA----- 1626  
QY 211 GlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAsp 230

```

Db 3091 ACATCCGACGATCATCATGTCACGACGAGAGTGAATCAAAACGT----- 3138
      |||::: |||::: |||:::
Qy 224 LysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaIleGlu 243
      |||::: |||::: |||:::
Db 3139 -----AAAAGTCGTCATTAAAGTCGAA 3159
      |||::: |||::: |||:::
Qy 244 GlnLysSerLysAspLeuGluGly-----ThreMetAspThrValAsnThr 258
      |||::: |||::: |||:::
Db 3160 TCGCAGACGATCGAGCTTTTCAGGCTCCACGAGTGCCTCCGACGCGATCATCAACG 3219
      |||::: |||::: |||:::
Qy 259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIleAlaIlePheThr 278
      |||::: |||::: |||:::
Db 3220 TCAACGAGTAGAGTACGTACAGACAGTACAGTACGATTAAGTGAATCGACAAAGTACA 3279
      |||::: |||::: |||:::
Qy 279 CysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIleAlaGlyAlaIleAlaIle 298
      |||::: |||::: |||:::
Db 3280 -----AGTCTTTCAGGCTCCACGAGTGC 3303
      |||::: |||::: |||:::
Qy 299 GlyAlaAlaGlyAlaIleAlaIleAlaIleThrValAlaIleThrGlnIleThrValGlnAlaVal 318
      |||::: |||::: |||:::
Db 3304 TCGACATCCGACAGTATCAACGTCACACGACGAGTGCATCAACAGTGAAGTACA 3363
      |||::: |||::: |||:::
Qy 319 ValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIle 338
      |||::: |||::: |||:::
Db 3364 -----TCGTTGAGTATCATTAAGCAGCAGCGTTTCAGATTCACAGACGCGCTCAAG 3417
      |||::: |||::: |||:::
Qy 339 ---LysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
      |||::: |||::: |||:::
Db 3418 TCGAAGACGTCCATCCACATCAACGAGTACAGTCAAAACAGTGCACACATCGTTA 3477
      |||::: |||::: |||:::
Qy 358 AlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377
      |||::: |||::: |||:::
Db 3478 AGTGCATCTTTAAGTACAAAGCATTTTCAGATTCACAGTACGTCGACAGTGC 3537
      |||::: |||::: |||:::
Qy 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397
      |||::: |||::: |||:::
Db 3538 TCACAGCTCAACGAGTACGAGTGCAGACAGTACACAGTACCTCA 3582
      |||::: |||::: |||:::
Qy 398 GlyValGlyValValAlaAlaIleProAlaLeuGlyLysGlyIleMetGlnMetGln 417
      |||::: |||::: |||:::
Db 3583 -----TTA 3585
      |||::: |||::: |||:::
Qy 418 SerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaIle 437
      |||::: |||::: |||:::
Db 3586 AGTGAATCGACACAGCACA-----AGTCTT 3609
      |||::: |||::: |||:::
Qy 438 AlaAspMetIleSerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLys 457
      |||::: |||::: |||:::
Db 3610 TCGGACTCAACAGTACGTCACAG-----TCAGAAAGTGCATCAACAG 3651
      |||::: |||::: |||:::
Qy 458 GlnThrGlyLysSerAsnGluMetThrGlnLysAlaThrLysLeu 472
      |||::: |||::: |||:::
Db 3652 TCAACGAGTGAGAGT---GACTCAACAGAAAGTACGTCATT 3693
      |||::: |||::: |||:::

```

```

; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
us-09-134-001C-1685

Alignment Scores:
Pred. No.: 3,87e-05 Length: 3561
Score: 150.50 Matches: 89
Percent Similarity: 37.47% Conservative: 98
Best Local Similarity: 17.84% Mismatches: 241
Query Match: 6.35% Indels: 71
DB: 4 Gaps: 14

us-09-889-314-2 (1-496) x us-09-134-001C-1685 (1-3561)

Qy 2 ThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21
      |||::: |||::: |||:::
Db 2210 ACGAGTACGTCGACATCCGACAGTCCACATCAACGTCACAGTACGATGACTACAGAGT 2151
      |||::: |||::: |||:::
Qy 22 ValLeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGluThr 41
      |||::: |||::: |||:::
Db 2150 GAAAGTACGTCATTAAAGCAGTCGATA-----ACACAAACGCTTCAGATTCAACAAGC 2097
      |||::: |||::: |||:::
Qy 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIle 61
      |||::: |||::: |||:::
Db 2096 GCGTCACATCCGACAGTCCGTCACATCAACGAGTGAAGTGAAGTGCAGAGTGCAGT 2037
      |||::: |||::: |||:::
Qy 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln 81
      |||::: |||::: |||:::
Db 2036 ACATCGTTAAGTGGCTCGACAAAGTACAGTCTTCAGATTCAACAGATTCACAGATCA 1977
      |||::: |||::: |||:::
Qy 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyValAspThrGly 101
      |||::: |||::: |||:::
Db 1976 GAC-----AGTCAATCAACGTCACACAGTGAAGAGGAGTCCGAAAGACAAAGTACATCC 1923
      |||::: |||::: |||:::
Qy 102 ValSerGlyAlaAlaIleAlaIleThr-----ThrAlaSerAsnThrAlaThrLys 116
      |||::: |||::: |||:::
Db 1922 TTAAGTGGCTCGACAAAGTACAGTCTTCAGATTCAACAGATTCACAGATCAACAGAT 1863
      |||::: |||::: |||:::
Qy 117 IleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGluSer 136
      |||::: |||::: |||:::
Db 1862 GCATCAACATCAACAGAGTGTGAGTACTCAACAGTGCAGACAGTCAATCAAGTGCCTCA 1803
      |||::: |||::: |||:::
Qy 137 Leu-----GlnSerLeuSerAlaAla 143
      |||::: |||::: |||:::
Db 1802 TTAAGTACAAAGCGTTTACAGCTCAACAGTACATCAACATCCGATGTCATCTGCATCA 1743
      |||::: |||::: |||:::
Qy 144 GlnMetLysGluValGlnAlaValAlaIleAlaLeuSerGlyLysSerSerGlySer 163
      |||::: |||::: |||:::
Db 1742 ACAAGTGAAGCAGCAGCTCGAAAGACAAAGTACATCGTTGAGTGCCTCAACAGTACAGC 1683
      |||::: |||::: |||:::
Qy 164 AlaLysLeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183
      |||::: |||::: |||:::
Db 1682 ATTTCAGACTCAACAGTACATCAACATCAAGAGTGCATCAACATCAACAGTGTGAGT 1623
      |||::: |||::: |||:::
Qy 184 Glu-----IleGlyLeuAlaLeuAlaLysAlaIleGlnThr---LeuGlyGlnAla 199
      |||::: |||::: |||:::
Db 1622 GAGTCAAAAGTACAGACAGCATCCATAGTGAATCATTAAGTACAGACGCTTCAGATCA 1563
      |||::: |||::: |||:::
Qy 200 ThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLys 219
      |||::: |||::: |||:::
Db 1562 ACGAGTACATCGACATCAACAGTGCAGACAGTCAACATCAACAGAGTGCAGTGCAGC--- 1506
      |||::: |||::: |||:::
Qy 220 LeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGlyThrGlnGluMet 239
      |||::: |||::: |||:::
Db 1506 ----- 1506
      |||::: |||::: |||:::
Qy 240 LysAlaAlaGluGlnLysSerLysAspLeuGlyThrMetAspThrValAsnThrVal 259
      |||::: |||::: |||:::
Db 1505 ---AGTCAAGTACGTCATCAACAGTACAGTACAGTACAGTACAGTACAGTACAGT 1449
      |||::: |||::: |||:::
Qy 260 MetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCys 279
      |||::: |||::: |||:::
Db 1448 ACATCCACATCAAGATTAAGTGCATCCATCAACAGAGTGAAGTGAAGTCAACAGTACA--- 1392
      |||::: |||::: |||:::

```

```
QY 145 MetLysGIuValGluAlaValAlaAlaLeuSerGlyLysSerSerGlySerAla 164
    |||
    |||
    |||
Db 2311 AGTAATTCAGACAGACAGAGCGTTTCACATTCACAAAGTACATCCAGACAGTGGC 2252
QY 165 LysLeuGIuThrProGluLeuProLysProGlyValThrProArgSerGluValIleGlu 184
    |||
    |||
    |||
Db 2251 TCCACATCAACGAGTGC-----AGTGCATCAGACAGT 2219
QY 185 IleGlyLeuAlaLeuAlaAlaIleGlnThr-----LeuGlyGluAlaThrLysSerAla 203
    |||
    |||
    |||
Db 2218 GCAAGTACGATCAATTAAAGATTCGACAAAGTACAAAGCGTTTCAGATTCAACGAGTGCAGTGC 2159
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
    |||
    |||
    |||
Db 2158 ACATCCGACAGTGCATCATATGTCAGCGACGAGAGTGAATCAAAAGT----- 2111
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGlnLysAlaAlaGlu 243
    |||
    |||
    |||
Db 2110 -----AAAAGTACGTCATTAAAGTGA 2090
QY 244 GlnLysSerLysAspLeuGlnGly-----ThMetAspThrValAsnThr 258
    |||
    |||
    |||
Db 2089 TCGACAACTACGAGCTTTCAGCGTCAACAGAGTGGCTGCATCCGACAGTGCATCAACG 2030
QY 259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThr 278
    |||
    |||
    |||
Db 2029 TCAACGATGACAGCTGACGACGACGATCAAGTACGTCATTAAAGTATCCGACAAAGTACG 1970
QY 279 CysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGly 298
    |||
    |||
    |||
Db 1969 -----AGTCTTTCAGCGTCAACGAGTGGC 1946
QY 299 GlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAlaVal 318
    |||
    |||
    |||
Db 1945 TCGACATCCGACAGCTCAACAGTCAACAGCGAGAGTCAACAGTCAACAGTCAACAGTCAAC 1886
QY 319 ValGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIle 338
    |||
    |||
    |||
Db 1885 -----TCGTTGAGTCAATTCATTAAAGCACAAGCGTTTCAGATTCAACAAAGCGCGTCAAG 1832
QY 339 ---LysAlaAlaValLysSerGlyLysLeuAlaPheIleLysThrLeuValLysAlaIle 357
    |||
    |||
    |||
Db 1831 TCAGAAATGCTCCACATCAACAGAGTGCAGTCAACAGAGTCAACAGAGTCAACAGTCAAC 1772
QY 358 AlaLysAlaIleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAla 377
    |||
    |||
    |||
Db 1771 AGTGAATCAATTAAAGTCAACAGCATTTTCAGATTCAACAGAGTCAACAGTCAACAGTCAAC 1712
QY 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397
    |||
    |||
    |||
Db 1711 TCAACGTCAACAGAGTGCAGTCAACAGTCAACAGTCAACAGTCAACAGTCAACAGTCAAC 1667
QY 398 GlyAlaGlyValAlaValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu 417
    |||
    |||
    |||
Db 1666 -----TTA 1664
QY 418 SerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAla 437
    |||
    |||
    |||
Db 1663 AGTGAATTCACACAGCA-----ACTCTT 1640
QY 438 AlaAspMetIleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLys 457
    |||
    |||
    |||
Db 1639 TCGGACTCAACAAGTACGTCAGC-----TCAGAAGTGCATCAACG 1598
QY 458 GlnThrGlyGluSerAsnGluMetThrGlnLysAlaThrLysLeu 472
    |||
    |||
    |||
Db 1597 TCAACGAGTGCAGAGT---GACTCAACAAGTGAAGTACGTCATTAA 1556
    |||
    |||
    |||
RESULT 34
US-09-134-001C-1626
; Sequence 1626, Application US/09134001C
; Patent No. 6380370
```

```
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1626
; LENGTH: 6414
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1626

Alignment Scores:
Pred. No.: 8,95e+05 Length: 6414
Score: 151.00 Matches: 93
Percent Similarity: 36.57% Conservative: 88
Best Local Similarity: 18.79% Mismatches: 200
Query Match: 6.37% Indels: 114
DB: Gaps: 14

US-09-889-314-2 (1-496) x US-09-134-001C-1626 (1-6414)

QY 5 SerLysSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr 24
    |||
    |||
    |||
Db 2470 AGGTAAAGCGATCAATCATTAATCAACAAAGTACAGTCTCTCAGAGTCTTAAGTGGC 2529
QY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIle 44
    |||
    |||
    |||
Db 2530 TCGAGTCGAATAGTACGTCATTCACAGCGACGAGTCAACGACGACGACGACGACGACGACGAC 2589
QY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
    |||
    |||
    |||
Db 2590 TCAAGATCAACGCG-----AGTACGTCGACATCTGATAGTCATCAGAGTCAACG 2637
QY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
    |||
    |||
    |||
Db 2638 AGAAAGAGTGAAGTCAACGAGTAAAGTACGTCATTAAAGTAAAGTCAACAGCAACAGCGTT 2697
QY 85 AlaAlaGlyLysGluSerSerGlnLysAlaGlyAlaAlaAspThrGlyLysSerGly 104
    |||
    |||
    |||
Db 2698 TCAAGATTCGACAAAGCGTTCACAGCTCAGAAAGTGCCTCAACATCAACAAAGTGTGAGTGC 2757
QY 105 AlaAlaAlaThrThrAlaSer---AsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
    |||
    |||
    |||
Db 2758 TCAACAAGTCAACAAGCATTTTCAGACTCAACGACACATCGACGTCGATGATGCGCATCAATA 2817
QY 124 GluGluAla----- 126
    |||
    |||
    |||
Db 2818 AAAGCGAGCGAATTCAGCAACAGCAAGTAAACTATTATCAAGATCAAGTGCAGTGCAGC 2877
QY 127 -----SerLysSerMetGlnSerThrLeuGlnLysLeuGlnSerLeuSerAlaAlaGln 144
    |||
    |||
    |||
Db 2878 TCTGATAGCGCATCAACGTCACAGGAGCGTCAAGTACGATCCAAATAGCCAAAGTACGTCATTA 2937
QY 145 MetLysGIuValGluAlaValAlaAlaLeuSerGlyLysSerSerGlySerAla 164
    |||
    |||
    |||
Db 2938 AGTAATTCAGACAGCAACAGCGTTTCAGATTCAACAAAGTCAACATCCGACAGTGC 2997
QY 165 LysLeuGlnThrProGluLeuProLysProGlyValThrProArgSerGluValIleGlu 184
    |||
    |||
    |||
Db 2998 TCCACATCAACGAGTGC-----AGTGCATCAGACAGT 3030
QY 185 IleGlyLeuAlaLeuAlaLysAlaIleGlnThr-----LeuGlyGluAlaThrLysSerAla 203
    |||
    |||
    |||
Db 3031 GCAAGTACGTCATTAAAGTGAATCGACAAAGTCAACGCGTTTCAGATTCAACAGAGTACGTCG 3090
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
```

Alignment Scores:	Pred. No.:	7.3e-05	Length:	6530
Score:		152.00	Matches:	79
Percent Similarity:		33.33%	Conservative:	28
Best Local Similarity:		24.61%	Mismatches:	162
Query Match:		6.42%	Indels:	52
DB:		5	Gaps:	7

US-09-889-314-2 (1-496) x PCT-US93-03993-1 (1-6530)
<p>59 AlaThrIleAlaIleAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThr</p> <p>4067 GCCACCGAGAGAGACCGCGCTCCGACGTGGAGCCACCGCTCCGCTCCGAGTAGCTCC</p> <p>79 AlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAla</p> <p>4007 TCCGCACACCACCGCGCGTGCCTCCGGAAGAGCACCCTCCGAGCC-----ACCGCGCT</p> <p>99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla</p> <p>3953 GCCCCCGAAGAGACCGCGCCGACACCGCCGCGCTCCGAGTAGACCTCCACCGCACC</p> <p>119 MetGlnThrSerIleGlnGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGln</p> <p>3893 ACCGCTGCTGCCGCCACACCGCTGGAGTAGACAGCCACCGCGCCACCGCCGACAGGA</p> <p>139 SerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaAlaLeuSerGly</p> <p>3833 GCCACCTCCACAGCTACACCTCTCCGGAAGAGCCCTCCGTGAGCTCTGTGGGGGC</p> <p>159 LysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysPro-----GlyVal</p> <p>3773 GCAGAGAGCTGACTGCTGCTGAGAGAGATATATATACCCCTCCGCCGCCAGAGAGCC</p> <p>177 ThrProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu</p> <p>3713 GCCCGCGAGCTGCACCGCCACCGACCGAATGACCGCGCGCGCTGCGCGCAACT</p> <p>197 GlyGluAlaThrLysSerAlaLeuSer-----AsnThrAlaSer</p> <p>3653 GGAGCCACACACAGAGAGCGCGCTCCGACGTAGAGCTCCGCCCGCGAGTATACCGCC</p> <p>210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIle</p> <p>3593 ACCGAGCTGAGGCTCCGCCCGCGAGTAGCGCGCGCGAGCTGAGGCC-----</p> <p>220 AspLysGluArgGluGluThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeu</p> <p>3543 -----</p> <p>250 GlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVal</p> <p>3542 -----ACCCTAGAACCGCTCCGTAGCTCTGGCACTGATTACTGTGGCA</p> <p>270 IleSerIleValAlaAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly</p> <p>3497 CGATCCGCCGCCGCCGCC-----GGAGGAACACACTCCGACACTCGTGTCC</p> <p>290 AlaAlaValAlaAlaAlaAlaAlaGlyValAlaAlaGlyValAlaAlaAlaIleThrVal</p> <p>3449 GCCGCCACCGGAATAGCGCGCGCGCGCGCGAGCTGATCAACCGCACCTGACTACCC</p> <p>310 AlaThrGlnIleThrValAlaGlnAlaValAlaGlnAlaValLysGlnAlaValAlaIleThrAla</p> <p>3389 GCCGCCACAGCTAGAGCGCGCGCTGCTTCACCGCC-----GCCGGAATACCGCC</p> <p>330 ValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPhe</p>

```

Db      3338  TCCGAGAGTGGAGGCCACCGCGCCACGACGAGTACTGTCCGCGTCGCGA----- 3291
Qy      350  IleystrhleuValysAlaIleAlaIlysaIalleSerlyGlyIleSerlyValPhe 369
Db      3290  -----ACCGCGCGCGCAGCCAGCGAGCTGCGCGCTCCAGAGAGCGCTCCGAGTAGGA 3237
Qy      370  Ala 370
      |||
Db      3236  GCC 3234

RESULT 33
US-09-134-001C-1515/c
: Sequence 1515, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lym Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GNC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 1515
: LENGTH: 2835
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1515

Alignment Scores:
Pred. No.:      2,38e-05      Length:      2835
Score:          151.00      Matches:      93
Percent Similarity: 36.57%      Conservative: 88
Best Local Similarity: 18.79%      Mismatches: 200
Query Match:      6.37%      Indels:      114
DB:               4      Gaps:      14

US-09-889-314-2 (1-496) x US-09-134-001C-1515 (1-2835)
Qy      5  SerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValLeuThr 24
      |||||
Db      2779  AGGTGTAAGGACGATCAATCATTCATCAACAATAAGAGCTCTCGAGAGCGTTAGTGGC 2720
Qy      25  SerThrProGlnGlyValProGlnGlnAspIlyLeuSerGlyAsnGluThrIlyScinIle 44
      |||||
Db      2719  TCGAGCTCAATATGATACGTCAATCCAGCGACGACGACGACGACCAACGACCAATTAACAATA 2660
Qy      45  GlnGlnThrArgGlnGlnIlyAsnThrGlnMetGlnIlySerAspAlaThrIleAlaGlyAla 64
      |||||
Db      2659  TCAGATATCGCG-----AGTACGTGACATCTGATAGTGCATCAAGATCAGCA 2612
Qy      65  SerGlyIlyAspIlySerSerSerThrIlyIlySerGlnIlyThrAlaProGlnGlnIlyVal 84
      |||||
Db      2611  AGAAAGACGTACGACGACGAGTAAAGTAGCTATTAAAGTAATGCAATCAACAAGTGTGAGTGGC 2552
Qy      85  AlaAlaGlyIlyLysIlySerSerGlnIlySerGlnIlyAlaGlyAlaAspThrGlyAlaSerGly 104
      |||||
Db      2551  TCGATTTCGACCAAGCGGTGTCACGTCAGAAAGTGGCTCAACATCAACAAGTGTGAGTGGC 2492
Qy      105  AlaAlaAlaThrThrThrAlaSer--AsnThrAlaThrIlyLysAlaMetGlnThrSerIle 123
      |||||
Db      2491  TCACAAAGTATCAACGATTTTCAGACTCAACGAGACATCGACGCTGATAGCGCATCAATA 2432
Qy      124  GlnGlnAla----- 126
      |||||
Db      2431  AAGGACGAGGAAATCAGCAAGCAACAAGTAATATTATCGAATCACTGAGTAGTGGAGC 2372
Qy      127  -----SerIlySerMetGlnIlySerThrLeuGlnSerLeuGlnIlySerLeuSerAlaIleGln 144
      |||||
Db      2371  TCTGATATGCGCATCAACGTCACAGCGATAGAGACTCCATATAGCCCAAGTACAGTCATTA 2312

```



```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,240

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/146,930

FILING DATE:
APPLICATION NUMBER: 07/876,286

FILING DATE: April 30, 1992
APPLICATION NUMBER: No. 6143727 yet assigned (204/132)

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 204/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-458-240-1

Alignment Scores:
Pred. No.: 7.3e-05 Length: 6530
Score: 152.00 Matches: 79
Percent Similarity: 33.33% Conserves: 28
Best Local Similarity: 24.61% Mismatches: 162
Query Match: 6.42% Indels: 52
Gaps: 7

US-09-889-314-2 (1-496) x US-08-458-240-1 (1-6530)
OY 59 AlAthrllleaIagIyAlaserGIyAspyLsrThrSerSerThrlThlyStrgluthr 78
||||| ||||| ||||| ||||| |||||
Db 4067 GCCACCAGAGAGCCCGCTCCGGACGTGGAGCACCGCCTCCGCCGTCCGGAGTACCTCC 4008

OY 79 AlaPrGInGlngIyAlaAlaIagIyLyGsUserSerGIuserGIuserglNlysAlaGIyAla 98
::: ||||| ||||| ||||| ||||| |||||
Db 4007 TCCGCACCAACCACCCCTGCTCCCGGAAGAGCCACTTCGCAGCC-----ACCGCGCT 3954

OY 99 AsphrclIyAlserGIyAlaAlaAlaIarThrrhlAlaserAstThAlaIThrLyslleAla 118
||| |::: ||| ||||| ||||| ||||| |||||
Db 3953 GCCCCGAGAGAACCCCGCCACACCGTGAGTAGCAGCACCGCCGCGCCGCCGAGAGA 3894

OY 119 MetGlnThrSerllleGluglAlaserLySerMetGlUserThrlleuGIuserleuGlN 138
::: ||| ||| ::: ||| ::: |||
Db 3893 ACCGCTCTGCCGCCACACACCGTGAGTAGCAGCACCGCCGCGCCGCCGAGAGA 3834

OY 139 SerLeuSerAlaAlaIagImetLySguIvaIgIuaIaValaIaAlaAlaIaIaIaIaIaIa 158
::: ||| ||| ||| ||||| ||||| ||||| |||||
Db 3833 GCCACCTCACAGCTACACACCTCTCCGGAAGACCCCTCCGTAGCTGTGGGGGC 3774

OY 159 LysSerSerelYserAlalylsleuGIutlrProGIuePrlyPro-----GlyAl 176
::: ||| ||| ||| ||| ||| ||| |||
Db 3773 GCACGAGCTGTACTGTCTGCTGAGAGAGATATAGCCCCCTCCGCCGAGAGAGCC 3714

OY 177 ThrProArgSerGIuallleGIuIlleGIyLeuAlaIaIaIaIaIaIaIaIaIaIaIaIa 196
::: ||| ||| ||| ||| ||| ||| |||
Db 3713 GCGCGCGCAGCTGCCACCGCCACACCGATATAGCCGCCGCCGCTGCGCGCCAACT 3654

OY 197 GlyGluAlaThrLysserAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 209
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3653 GGACGACCAACAAGAGAGCCGCTCCGAGCTAGAGCTCCGCCGCCGAGTACCGCC 3594

OY 210 ThrlnaIaIaIaIaspGlnThrAsnLysleuGIyLeuGIuserglNlysAlaIaIaIaIaIa 229
||| ||| ||| ||| ||| ||| ||| |||
Db 3593 ACCGCTGTGAGCTCCGCCGCCGAGTAGCCGCCGCCGACGTGAGCC----- 3543

```

```

Oy      230 AsplysGlnArgGluGluTyrGlnGluMetLysAlaAlaGluInLysSerLysAspLeu    249
Db      3543 -----                               3543
Oy      250 GluGlyThrmelAspThrValasnrThrValMelileAlaIvalSerValAlaIleThrVal    269
Db      3542 -----ACCGGTAAACCAGCCTCCGTAAGTCCTGGCACTGATACTGTGGGA    34988
Oy      270 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaIcy    289
Db      3497 CGAATCCCCCGCCGCCAC-----GGAGAACCACCTCCGCAGACTCTGTCCC    34500
Oy      290 AlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaIleThrVal    309
Db      3449 GCCCGCACCGGAATAATGCGCGCGCGCGCGCGCGCGCATTCACCGCACCTGACTACC    33900
Oy      310 AlAtHrgInIleThrValGlnAlaValGlnAlaValGlnAlaValGlnAlaValIleThrAla    329
Db      3389 GCGCGCACAGCTAAGAGCGCGCGCGCTGCCTCCACCGC-----GCCGGAATTACCGCC    33330
Oy      330 ValArGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSeryIlyLysAlaPhe    349
Db      3338 TCCCAGCTGGAGGACCGCGCGCGCACAGAGTACTTGCCGCGCTCCGGA-----    32910
Oy      350 IleYsrThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPhe    369
Db      3290 -----ACCGCGCGCGCGCACGACCGACCGAGCTGCCGCTCCAGAGAGACTCCGCACTAGGA    32370
Oy      370 Ala    370
Db      3236 GCC    3234

RESULT 32
PCT-US93-03993-1/c
; Sequence 1, Application PC/RUS9303993
; GENERAL INFORMATION:
;   APPLICANT: Roop, Dennis R.
;   APPLICANT: Rothnagel, Joseph A.
;   APPLICANT: Greenhalgh, David A.
;   TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL
;   TITLE OF INVENTION: VECTOR SYSTEMS
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Fulbright & Jaworski
;     STREET: 1301 McKinney, Suite 5100
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: U.S.A.
;     ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US93/03993
;   FILING DATE: 19930428
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5325
; TELEFAX: 713/651-5246
; TELEX: 762829
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6530 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear

```

```

1  APPLICANT: Rothnagel, Joseph A.
2  APPLICANT: Greenhalgh, David A.
3  TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
4  TITLE OF INVENTION: AND METHODS OF USE
5  NUMBER OF SEQUENCES: 4
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: LYON & LYON
8  STREET: 611 West Sixth Street
9  CITY: Los Angeles
10 STATE: California
11 COUNTRY: U.S.A.
12 ZIP: 90017
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
16
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
19 SOFTWARE: WordPerfect (Version 5.1)
20
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/146,930
23
24 FILING DATE:
25 CLASSIFICATION: 435
26
27 PRIOR APPLICATION DATA: Including application two
28 PRIOR APPLICATION DATA: described below:
29 APPLICATION NUMBER: 07/876,286
30 FILING DATE: April 30, 1992
31 APPLICATION NUMBER: No. 5958764 yet assigned (204/132)
32 FILING DATE: October 29, 1993
33
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Warburg, Richard J.
36 REGISTRATION NUMBER: 32,327
37 REFERENCE/DOCKET NUMBER: 204/152
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: (213) 489-1600
40 TELEFAX: (213) 955-0440
41
42 TELEX: 67-3510
43
44 INFORMATION FOR SEQ ID NO: 1:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 6530 bases
47 TYPE: nucleic acid
48 STRANDEDNESS: single
49 TOPOLOGY: linear
50 MOLECULE TYPE: DNA (genomic)
51
52 US-08-146-930-1
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
9
```

Db	3883	GCCACCGTCCACAGCTACGACTCTCCGGAGAGACCCCTCTGACTCTGCTGGGGGGG	3774
0y	159	LysSerSerGlySerAlaIleuLysLeuGlnThrProGluLeuProIlyPro-----GlyVal	176
Db	3773	GCAGGAGCTGTACTGCTGCTGAGAGAGAGATAATAGCCCTTCGCCGCCACAGAGGAGCC	3714
0y	177	ThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaIleuAlaIleGlnThrLeu	196
Db	3713	GCCGCCACACTGCCACCGCCACACACCGAGATAGCCGCCGCCGCTGCGCGCGACT	3654
0y	197	GlyGluAlaThrTrpLysSerAlaLeuSer-----AsnTrpAlaSer	209
Db	3653	GGAGCCACACACAGAGAGAGCGGCTCCGACGTAGAGCTCCGCCCCGAGATAGCGCGCC	3594
0y	210	ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIle	229
Db	3593	ACCGAGCTGGAGGCTCCGCCGCCCGGAGTAGCGCCGCCGCCACGTGGAGCC-----	3543
0y	230	AspLysGluArgGluGlnIuTrpGlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeu	249
Db	3543	-----	3543
0y	250	GluGlyThrMetAspTrpValAsnThrValMetIleAlaValSerValAlaIleThrVal	269
Db	3542	-----ACCGGTAGAACCGCTCCGTGAGCTGTGGCACTATACTGTGGGA	3498
0y	270	IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly	289
Db	3497	CGATTCGGCGCGCGCCGACCC-----GGAGAACACACTCCGCGAGCTCTGACC	3450
0y	290	AlaAlaValaGlyAlaAlaAlaAlaGlyGlyValaAlaGlyAlaAlaAlaAlaThrThrVal	309
Db	3449	GGCGGACCGGAATAGCCGCGCGCGCGCGCGACCTGGATCCAGCCGACCTGAGTACC	3390
0y	310	AlaThrGlnIleThrValaGlnAlaValaValaGlnAlaValaLysGlnAlaValIleThrAla	329
Db	3389	GCGCGCACAGCTAGAGCCGCCGCTCTCCACCGCC-----GCCGGAATACCGCC	3339
0y	330	ValArgGlnAlaIleThrAlaIleLysAlaIleLysSerGlyIleLysAlaPhe	349
Db	3338	TCCGACACTGAGCCACGCCGCCGCCACAGAGTACTGCCGCTTCGG-----	3291
0y	350	IleLysThrIleuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPhe	369
Db	3290	-----ACCGCGCGCGGACGACACCGAGCTGCGCGCTTCAGAGAGGCTCCGACGTAGGA	3237
0y	370	Ala 370	
Db	3236	GCC 3234	
<p>RESULT 31</p> <p>US-08-458-240-1/C</p> <p>Sequence 1, Application US/08458240</p> <p>Patent No. 6143727</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Roop, Dennis R.</p> <p>APPLICANT: Rothnagel, Joseph A.</p> <p>APPLICANT: Greenhalgh, David A.</p> <p>TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS</p> <p>TITLE OF INVENTION: AND METHODS OF USE</p> <p>NUMBER OF SEQUENCES: 4</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: LYON &amp; LYON</p> <p>STREET: 611 West Sixth Street</p> <p>CITY: Los Angeles</p> <p>STATE: California</p> <p>COUNTRY: U.S.A.</p> <p>ZIP: 90017</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: 3.5" diskette, 1.44 Mb storage</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: IBM MS-DOS (Version 5.0)</p> <p>SOFTWARE: Wordperfect (Version 5.1)</p>			

```

RESULT 29
US-08-591-079-5
; Sequence 5, Application US/08591079
; Patent No. 5972899
; GENERAL INFORMATION:
; APPLICANT: Zychlinksky, Arturo
; APPLICANT: Chen, Yajing
; TITLE OF INVENTION: Apoptosis Induced by Shigella IpaB
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,079
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ljvat, Shmuel
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: 15661-20017.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0764
; TELEX: 90-4030 MRSNFORSMWH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3106 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: yersinia pseudotuberculosis
; STRAIN: ypIII(PIBI)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12..1217
; OTHER INFORMATION: /gene= "yopB"
US-08-591-079-5

Alignment Scores:
Pred. No.:      1,95e-05      length:      3106
Score:          152.50       Matches:     93
Percent Similarity: 37.53%   Conservative: 56
Best Local Similarity: 23.43% Mismatches:    163
Query Match:      6.44%     Indels:        85
DB:                2         Gaps:         17

US-09-889-314-2 (1-496) x US-08-591-079-5 (1-3106)
OY      152 ValValAlaAlaLeuSerGlyLySerserGlySerAlaLysLeuGlutProGluLeu 171
           |||      |||:::  |||      ::      ::|||      :|||
Db      111 GTGGCGGGAGAACTGAAGAGATAAATGGCGGGGTGACTTCT--CAGGCCTACACCTC 167
OY      172 ProGlyProGlyValThrProArgserGluValIleGu---IleGlyLeuAlaLeuAla 190
           |||      |||      |||:::  |||      |||      ::|||      :|||
Db      168 CTGCACCACCTAGCAGCTGGTTGCCAGCCCAAGTACTGAAGACACAACAGCAAGAAGTCACT 227
OY      191 LysAlaIleGln-----ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyraLa 208
           |||      :::::  |||      |||      |||      ::|||      |||
Db      228 AATATTATTGGAGTGCGTCAACCCGGCGGCGGCGCAGAGACTCAACTGATCATTAATTAATGTT 287
OY      209 Ser-----ThrGlnAlaGlnAlaAsp-----GlnThrAsnLys 219

```

```

Db      288 TCAATGCTAACGAAGTTTACGCTGCTTACCGTATACATTGGAGATTGAGTATAGTAAG      347
OY      220 Leu-----GlyLeuGlu-----LysGlnAlaIleLysIle--AspLysGluArgGlu      234
Db      348 CTAGCTTCTTAATTAGAAAGAGTACGCAAGCAAGATATAAAATCGCTGATTTTCACGCTCTT      407
OY      235 GluTyrGlnGlnuMetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyTThMetAsp      254
Db      408 CATGACAAACAACATGAGAAATTTGAAAGAAATCAAGAGAAATATCAAGAAACAGAAAG      467
OY      255 ThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAla      274
Db      468 AATGCCAAGCAAGTCAAGAAATCC-----GGCATGGCATCA      503
OY      275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAla      294
Db      504 AAGATTTTTCGTGGCTGCACGCCCATAGGCTCATGATTTGCTGGCTCATGTGGTGC      563
OY      295 AlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaIleThrValAlaIleThrGlnIleThr      314
Db      564 TCAGGGGTAGAGAGCCGTTCGGCGGTGAATGATGATGCTTCAGGCGTAATGGCATGGCG      623
OY      315 ValGlnAlaValAlaGlnAlaValLysGlnAla-----      325
Db      624 AATATGGCAGTGAACAACAGCGCGGGAAGATGGCTGTATTTCCCAAGAGGCATGAATA      683
OY      326 -----ValIleThrAlaValAlaArgGlnAlaIleThr-----      335
Db      684 TTAGAGCGCGATACACTGCTGGCATTTGAAGTCGCAATTGACTGATGTTTCAACCGTATGACC      743
OY      336 -----AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThr      352
Db      744 TTTGGCGGTATCGGCACCTAAATATCTCGGTAAATATTTGGCGCAAAATCGGTCTAACCC      803
OY      353 LeuValLysAlaIleAlaLysAla-----IleSerLysGly      364
Db      804 GCAAGCTCTTGCGGCTAAAGAGACCGAGATTTCGGCCAAAGTTGCCCAATTTTCGACAGGC      863
OY      365 IleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSer      384
Db      864 ATATCAAAACACTGTGCGGAGTGCAGAGTAAATTTTAGGGCGAGTTTTCGTGTTAACAA      923
OY      385 LysValIleSerSerLeuThrSerLysIleThr-----ValGlyValGlyValVal      403
Db      924 ATGAGCCATGCAATCCGTACAGATCACAGGCAACACAACTGCGCGTGTGCTG-----      977
OY      404 AlaAlaProAlaLeuGlyLysGlyIleMetGln-----MetGlnLeuSerGlu      419
Db      978 -----GGCAGCGGAATTAATCATAGACCATATAATTAAGCAACACTGAT      1022
OY      420 MetGlnGlnAsnValAlaGlnPhe-----GlnLysGluValGlyLysLeuGlnAla      436
Db      1023 TTACACACTATTAACGCTGATTTTGGCTTGAACAAGCAGACATGGCAGCGTTAACAAAGT      1082
OY      437 AlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaSerLysIleAlaSer      456
Db      1083 ATTATTATGCCACATC-----AAA      1100
OY      457 LysGlnThrGlyGluSerAsnGluMetThrGlnLysAlaIleThrLysLeuGlyAlaGlnIle      476
Db      1101 GAAGGTATTCCTCATTTGTGAGATCCATCAACAAGTATGAGAACTGATTTTCCAGATG      1160
OY      477 LeuLysAlaTyrAlaAlaIleSerGlyAlaIleAlaGlyAlaHisLysThr      493
Db      1161 ATTATATGCAAAAGGTGACATGCTGCATTAATTTGGCCGCGAGCAACCCCATCT      1211

RESULT 30
US-08-146-930-1/C
; Sequence 1, Application US/08146930
; Patent No. 5958764
; GENERAL INFORMATION:
; APPLICANT: Koop, Dennis R.

```



```

OY 309 lAla-----ThngInIeThrValGlnAlaValAlaGlnAlaValysGI 324
    |||          |||          :|||          |||          :|||
Db 875 TGCGGTGAGTCACATCAACATGACTGACCTCACCAGCACATCATGACTTGTGAATC 934
                                     :|||          :|||
OY 324 nAlaValleThrAlaValAArgGlnAlaAlIethrAlaAlaIelysAlaAlaValysSe 344
    |||          :|||          |||          |||          :|||
Db 935 TGCATCAACCAAGTGCGTCAAGCTCAGCATTCGACAAAGCCG---TCAGTTCAAGCAAGT 991
                                     :|||          :|||
OY 344 rGlyllellysAlaPheIlellysThrLeuVallysaIleAlaIalylsAlaIleSerLysGI 364
    :|||          |||          |||          |||          :|||
Db 992 CAGTGGTCGTCAGCCTCAGCAAGTACCAAGTTCAGCTTCACGCGTCGACAAAGTCGTCGGG 1051
    :|||          |||          |||          |||          :|||
OY 364 y 364
    |
Db 1052 c 1052

RESULT 27
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:      1.25      Length:      4411529
Score:          155.50     Matches:       126
Percent Similarity: 31.37% Conservative:   44
Best Local Similarity: 23.25% Mismatches:    193
Query Match:      6.56% Indels:         180
DB:               Gaps:           18

US-09-889-314-2 (1-496) x US-09-103-840A-1 (1-4411529)
OY 59 AlArthrIleAlaGlyAlaSerGlyLysAspLysrThrSerSerThrThrLysrThrGlutThr 78
    |||          :|||          |||          |||          :|||
Db 3943780 GCCGCCCGTCCGGCTGCC-----GTTCGGCTT 3943754
                                     :|||          :|||
OY 79 AlaProGlnGlnGlyValAlaAlaIlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
    |||          :|||          |||          |||          :|||
Db 3943753 GCCGCCCGTCCGGCGGTC--GCCGGCGTCGCCGCACGCCCGCTTGGCCGCGGTGCC 3943697
    |||          :|||          |||          |||          :|||
OY 99 AspThrGlyValSerGlyAlaAlaAlaIlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
    |||          :|||          |||          |||          :|||
Db 3943696 GCCGCCCGTCCGCCACAGCGCCGCCGCCACCGTTACACACACACACCGTTCGCCGCC 3943637
    |||          :|||          |||          |||          :|||
OY 119 MetGlnTrpSerIleGluGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeuGln 138
    |||          :|||          |||          |||          :|||
Db 3943636 -----ACCGCGCGCGCGTGCATTGCC 3943618
    |||          :|||          |||          |||          :|||
OY 139 SerLeuSerAlaIlaGlnMetLysGluValGlu-----AlaValAlaVal 153
    |||          :|||          |||          |||          :|||
Db 3943615 GGAAATTACGGTTGCCGCCGCTTAACGCCCATTCGCTGTGCCGCCGCGTCCGCC 3943556
    |||          :|||          |||          |||          :|||
OY 154 AlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGluThr-ProGluLeuProLy 173
    |||          :|||          |||          |||          :|||
Db 3943555 GCCCGTGGCGCGTCGCTCCACCGCTGCCGCCGCCGCCACACCGCCAGCGTTAACGCC 3943496

```

QY 173 s-----ProglyalThrProArgSerGluValIleGI 184  
 Db 3943495 TTGTCCACCGGAGCGGCGGGAGCCGCGCCACCGCCGACGGCGGTGGCGCCGCCCGCGCC 3943436  
 QY 184 uileGlyLeuAlaLeuAlaLys-----AlaIleGlnThrLeuGlyGluAlaThrLysSerAl 203  
 Db 3943435 -GCCGGCAGCGCGCTGCCCGCCCGCCGCCACCCCTACCCCGCGTCCCGCCACCGGTGCCG 3943377  
 QY 203 aLeuSerAnfYrAlaSerThrGlnAlaGlnAlaAspGlnThrAnfLysLeuGI 223  
 Db 3943376 CTTGGCGCGCGCGCACACCTTCCTCGCGCGCGCGCGCGGTACCGG----- 3943331  
 QY 223 uLysGlnAlaIleLysIleAspLysGluArgGluGlnIuTyrgInGlnMetLysAlaAlaGI 243  
 Db 3943330 -----GGCCCCCGCGGTGCCG 3943314  
 QY 243 uGlnLysSerLysAspLeu-----GlnGlyThrMetAspThrVa 256  
 Db 3943313 CGGTGCCCGCGGTATATATGTGCCCGCGCTGCCCGCGCGCCGACACCAC 3943254  
 QY 256 lAsnThrValMetLysAlaValSerValAlaIleThrValIleSerIleValAla----- 274  
 Db 3943253 CTTGACCAACCGGACCGCGCGCGCGCGCGCGCTGGACCACTTGCACCGCGCGCGCGG 3943194  
 QY 275 -----AlaIlePheThrCysGlyAlaGlyLeuAlaGI 288  
 Db 3943193 CCCCCCGGTGCCCGCGCGCGCTCGGTGGCGCGCGCGCGCGCGCTCGCGCGCGCTC 3943134  
 QY 288 aGlyAlaAlaValaGlyAlaAlaAlaAla-----GlyGlyAlaAlaGlyAlaGI 304  
 Db 3943133 CGGTGCCGCGCGCGCGCGCGCGCGCTGGTGAAGAGAGCTGGCGCGCGCGCG 3943074  
 QY 304 aAlaAla----- 306  
 Db 3943073 CGGTGCCGCGCGCACCGCGCTTGCGCGCGGTGCCAGCGGTGCTTACCGCGCACCGG 3943014  
 QY 307 -----ThrThyValAlaThrGlnIleThrValGlnAl 317  
 Db 3943013 CGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGACCTTGTGCCCGG 3942954  
 QY 317 aValVal-----GlnAlaValLysGlnAlaValIleThrAlaValArg----- 331  
 Db 3942953 CGTTGCCCGCGGAGCTGTTGGGCGCGCTGCCCGCGCGCGCGCGCGCGCGCGACTT 3942894  
 QY 332 -----GlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPh 349  
 Db 3942893 CGCGGTCCGCGCGGTGGCGCGCTTGGCGCGCGCGCGCGCGCGACGATACGTCGCGCGG 3942834  
 QY 349 eIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGly----- 364  
 Db 3942833 CGCCACGACTCGCGCTTGGCGCGCGGTGGCGCGCGCGCGCGCGCGCGCGG 3942774  
 QY 365 -----I 365  
 Db 3942773 CCCCACGCTTCTGCGACGCTGGCGCGCGCGCGCGCGCGCGCGCTTGCCTGCCGCGCGC 3942714  
 QY 365 eSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPe----- 380  
 Db 3942713 CGGTACGCTTGGCGCGCGCGCGCGCGGTGTGTGCGCGCGCGCGCGCGCGCGCGGT 3942654  
 QY 381 -----ProLysLeuSerLysValIleSerSerLeuThrSerLysTr 394  
 Db 3942653 TTCCGCGCTTCCGCAATCCCGACCGCTTCCGCGCGGTATTTCCGCTACCGCGCGCTC 3942594  
 QY 394 pValThrValGlyValaGlyValaValAlaAlaProAlaLeuGlyLysGlyIleMetGI 414  
 Db 3942593 CGCCACG 3942534  
 QY 414 mMetGlnLeuSerGluMetGlnGlnAsnValaGlnPheGlnLysGluValGlyLysLe 434  
 Db 3942533 -----GGTACTC 3942484  
 QY 434 uGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrPglGlnAlaSerLysI 454

```

Db 397 GGTTCAGCATCAACGAGTAGC---TCAGCTTCAGGCGTCAACCAAGCTTCAGCTCAGCA 453
Oy 217 ThAsnLysLeuGlyLeuGlyLysGlnAlaIleLysIleAspLysGluArgGluLysTyr 236
Db 454 TCACACAGTGGTCAGGCTCA----- 474
Oy 237 GlnGluMetLysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal 256
Db 475 GCAAGTTCACAGGTCGTGAATCGCATCAACGATGCGCTGAGTCAACCATCAACGAGT 534
Oy 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 535 AGCTACGCTTCAGCAACGCTCAGCTCT------GAATCGGCC 573
Oy 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAla 296
Db 574 TCACACAGTGGTCAGGCTCAGCATCAACGATGCGCTCAGCTTCAGCAAGTACAGTGTCT 633
Oy 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 316
Db 634 TCAGGCTCAGGTCGACAGTGCCTGCGCTCAACGATGATGCTGAA-----TCG 684
Oy 317 AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaValArgGlnAlaIleThrAla 336
Db 685 GCATCAACCAAGTGGCTCAGGCTCAGCATCAACGATGATGCTGCTGCTCAGCATCAACCAAGT 744
Oy 337 AlaIleLysAlaAlaValLysSer 344
Db 745 GCCTGCGCTTCAGGCTCAACCAAGT 768

```

## RESULT 26

US-08-961-527-373  
Sequence 373, Application US/08961527  
Patent No. 6420135

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8512

TELEFAX: (301) 309-8504

INFORMATION FOR SEQ ID NO: 373:

SEQUENCE CHARACTERISTICS:

LENGTH: 1068 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-373

## Alignment Scores:

Pred. No.: 1.21e-06 Length: 1068  
Score: 157.00 Matches: 89  
Percent Similarity: 38.58% Conservative: 58  
Best Local Similarity: 23.36% Mismatches: 157  
Query Match: 6.63% Indels: 78  
DB: 4 Gaps: 10

US-09-889-314-2 (1-496) x US-08-961-527-373 (1-1068)

```

Oy 2 ThrAsnMetSerIleLeuSerSerSerGlyProAspAsnGlnLysAsnIle----- 18
Db 90 ACAGAGTGGCTGTGATCAGCATCAACGATGCGCTCAGCTCAGCAAGTACAGTCTTG 149
Oy 19 -----MetSerIleValLeuThrSerThrProGlnGlyValProGlnLysAspLysLe 36
Db 150 AATCTGATCAACAGTGGCTGACCTCAGCATGACAGAGGCGCTCAGCTCAGCAAGTAC 209
Oy 36 uSerClyAsn---GluThrLysGlnIleGlnIleThrArgGln----- 49
Db 210 CAGTGGCTCAGCTCAGGCTGACAAAGTGCCTGCGCTTCAGCAAGTACCAAGTCCGCTC 269
Oy 50 -GlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyLysAspLys 69
Db 270 CAGCAAGTACCAAGTGCCTGACCTCAGCTGACAGTGCCTGCGCTTCAGCAAGTAC 323
Oy 69 sThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaGlyLysG 89
Db 324 TGCATCTGAATGGCATCAACCAAGTGCCTCAGCTCAGCAAGT-ACATGATCAAGCTTC 382
Oy 89 uSerSerGluSerGlnLysAlaGlyAlaAspThrGlyLysSerGlyLysAlaAlaThrTh 109
Db 383 AGCATCAACAGTGCATCGCTCAGCATCAACGATGCGCTCAGCATCAAGTACCAAGTAC 442
Oy 109 rAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGluLysLysSe 129
Db 443 TGGCTGAGCTTCGGATCAACAAAGTGCCTGCGCTCAGCAAGTACTAGCGCTCAGCTC 502
Oy 129 rMetGluSerThrLeuGlnSerLeuGlnSerLeuSerIleAlaGlnMetLysGluValG 149
Db 503 AGGCTCAACAAAGTTCATCGCTCAGCTCAGCTCAGCAAGTGCCTG-----GCCTC 550
Oy 149 uAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlyLysAlaLysLeuGluThrPr 169
Db 551 AGCAAGTATCTCAGCGCTCAATCGCATCAACAAAGTGCCTGCGCTCAGCATCAACAG 610
Oy 169 oGluLeuProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAla 189
Db 611 T-----GCATC 616
Oy 189 uAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnThrAlaSe 209
Db 617 AGCTCAGCAAGCAACCGATGCGCTGCGCTCAGCAAGCAAGCGGCTGTAATCCGCATC 676
Oy 209 rThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysI 229
Db 677 AACCAAGTGCCTCAGCTCAGCAAGTAC----- 704
Oy 229 eaSpLysGluArgGluGluGlyGlnGluMetLysAlaAlaGlnLysSerLysAspLe 249
Db 705 -----TCAGCATGTAATCAGCATCAACCAAGT-- 731
Oy 249 uGluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVa 269
Db 732 -----GCATCGGCTTCAGCAAGCAAG 754
Oy 269 IleSerIleValAlaAlaIlePheThrCysGlyValAlaGlyLeuAlaGlyLeuAlaAla 289
Db 755 TGCTTCAGGCTCAGCAAGTATCTCAGCGTGAATCGGCATCAACGATGCTCGCTTC 814
Oy 289 yAlaAlaValAlaGlyAlaAlaAlaGlyGlyLysAlaGlyAlaAlaAlaAlaThrThrVa 309
Db 815 AGCAAGTACTAGCGCTCAGCATCAACGATGCGCTTCAGCTCAGCTCAGCTCAGCAAG 874

```

[illegible]

```

1 ADDRESS: Human Genome Sciences, Inc.
2 STREET: 9410 Key West Avenue
3 CITY: Rockville
4 STATE: Maryland
5 COUNTRY: USA
6 ZIP: 20850
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
10 COMPUTER: HP Vectra 486/33
11 OPERATING SYSTEM: MSDOS version 6.2
12 SOFTWARE: ASCII Text
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/961,527
15 FILING DATE:
16 CLASSIFICATION: 424
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER:
19 FILING DATE:
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Brookes, A. Anders
22 REGISTRATION NUMBER: 36,373
23 REFERENCE/DOCKET NUMBER: PB340P1
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (301) 309-8504
26 TELEFAX: (301) 309-8512
27 INFORMATION FOR SEQ ID NO: 372:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 807 base pairs
30 TYPE: nucleic acid
31 STRANDEDNESS: double
32 TOPOLOGY: linear
33
34 US-08-961-527-372
35
36 Alignment Scores:
37 Pred. No.: 5,41e-07 Length: 807
38 Score: 158.50 Matches: 66
39 Percent Similarity: 38.54% Conservative: 45
40 Best Local Similarity: 22.92% Mismatches: 140
41 Query Match: 6.69% Indels: 37
42 DB: 4 Gaps: 5
43
44 US-08-989-314-2 (1-496) x US-08-961-527-372 (1-807)
45
46 QY 57 SeraspalarhrllleaglyalaSerglyLysAsplysThrSerSerthrhrLysThr 76
47 ||| ::||| ||| ||||| ||| ||| ||| |||
48 Db 16 TCACGATCAACAAGTGGCTGGCTTCAGCATCAACAGTGCTGGCTTCAGCATCAAC 75
49 QY 77 GluthralaPrroglnglnglyValAlAlAlaglyLysLuserSerGluserGlnLysAla 96
50 ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
51 Db 76 AGTGGCTACCAATTCAGCAAGTACCAAGTGGCTTCAGTCACTCAACATCAACAAAGTCTTCAGCC 135
52 QY 97 GlyAlaAspThrGlyValSerGlyValAlAlAlaAlaThrThrAlaSerAsnThrAlaThrLys 116
53 ||| ||| ||| ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
54 Db 136 TCACGATCGACAAGTGGCTGGCTTCAGCAAGCAACATCAAGATCTGAATCAGGCTCAAC 195
55 QY 117 lleAlaMetGlnThrSerlleGluGluAlaSerLysSerMetGlnSerThrLeuGlnSer 136
56 ||| ::||| ::||| ||| ||| ||| ||| ||| |||
57 Db 196 AGTGGTGGCTTCAGCAAGTACCAAGTGGCTTCAGCTTCAGCATCAACACAGCGCTGGCC 255
58 QY 137 LeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnValAlaValAlaAlaLau 156
59 ||| ||| ||| ||| ::||| ||| ||| ||| |||
60 Db 256 TCACCAACACACCTCAGCTTCGAATCGAGCTCAACACAGCGCTTCGCGCTCAAGCACC 315
61 QY 157 SerGlyLysSerSerGlySerAlaLysLauGlnLthrProGlnLueuProLysProGlyVal 176
62 ||| ||| ||| ||| ||| ||| ||| ||| |||
63 Db 316 TCAGCTTCGAATCGGCTCA----- 336
64 QY 177 ThrProArgSerGlnValAlleGlnlleGlyLeuAlaLauAlaLysAlaIleGlnThrLau 196
65 ||| ||| ::||| ::||| ::||| ::|||
66 Db 337 ACCAGCGCTCAGGCTCAGCATCAACAGTGGCTTCAGCAAGCAACAGCGCTCG 366
67 QY 197 GlyGluAlaThrLysSerAlaLauSerAsnTyrLaseThrGlnAlaGlnAlaAspGln 216
68 ||| ||| ::||| ||| ||| ||| ||| ||| |||

```

```

OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 941 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-368

Alignment Scores:
Pred. No.: 1,08e-07 Length: 941
Score: 166.50 Matches: 70
Percent Similarity: 39.94% Conservative: 57
Best Local Similarity: 22.01% Mismatches: 180
Query Match: 7.03% Indels: 12
DB: Gaps: 2

US-09-889-314-2 (1-496) x US-08-961-527-368 (1-941)
QY 27 PROGLINGLYVALPROGLINGLYVALSERGLYASNGIUTHRLYSGINLEGLINGN 46
DB 20 CCTCACACAGTGGCTTCAGCCTCAGCCTCAACCAAGTGGCTTCCAGCAGAGTCACTG 79
QY 47 THRARGGLINGLYVALSERGLYASNGIUTHRLYSGINLEGLINGN 66
DB 80 CGTCACTTCAGCAGCAGCAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCCAGC 139
QY 67 LYSASPLYSRTHSERRTHRTHLYSTHGLUTHRALAPROGLINGLYVALAALA 86
DB 140 TCACACAGTGGCTTCAGCAGTCAACGAGTGGCTTCACTTCAGCAGTCAACCAAGT 198
QY 87 GLYLYSERGLYASNGIUTHRLYSGINLEGLINGLYVALASPTHRGLYVALSERGLY 106
DB 199 TCACCATCAGCATCAACCAAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCA 258
QY 107 ALATHRTHALASERASPTHRALATHRTHLYSLEALAMETGINTHRSERILEGLING 126
DB 259 TCACACAGTGGCTTCAGCAGTCAACGAGTGGCTTCACTTCAGCAGTCAACCAAGT 318
QY 127 SERLYSERMETGLUTRTHLEUGLUSERLEUGLUSERLEUGLUSERLEUGLUSER 146
DB 319 TCAGTCTCAGCAGCAGCAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCA 378
QY 147 GLUVALGLUVALAVALAVALAVALAVALAVALAVALAVALAVALAVALAVALA 166
DB 379 ACTATCTCAGGCTGTCGATTCGCAATCAACGAGTGGCTTCACTTCAGTCA 438
QY 167 GLUTRPROGLINGLYVALSERGLYVALTHRPROARGSERGLYVALILEGLILEGL 186
DB 439 TCGGCTTCAGCAGCAGCAGTGGCTTCACTTCAGCAGTCAACCAAGT 483
QY 187 LEUVALLEUVALAVALAVALAVALAVALAVALAVALAVALAVALAVALAVALA 206
DB 484 ---GCTTCAGCCTCAGCAGAGTATCTCAGCCTTGAATGGCATCAACCAAGTGGCT 540
QY 207 TYRALASERTHRGLINALAVALAVALAVALAVALAVALAVALAVALAVALA 226
DB 541 TCACGACAGTACTAGTGCATCAGCAGTCAACCAAGTGGCTTCACTTCAGTCA 585

```

```

QY 227 ILEYSILEASPLYSGLUARGGLINGLYVALSERGLYASNGIUTHRLYSGINLEGL 246
DB 586 GCTTCAGCAGTCAACCAAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCA 645
QY 247 LYSASPLYSGLUGLYTHRTHLYSTHGLUTHRALAPROGLINGLYVALAALA 266
DB 646 ACCAGCGCTTCAGCAGCAGCAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCT 705
QY 267 ILETHRVALILESERILEVALAVALAVALAVALAVALAVALAVALAVALAVAL 286
DB 706 GCTTCAGCAGTCAACCAAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCA 765
QY 287 ALAVALGLYVALAVALAVALAVALAVALAVALAVALAVALAVALAVALAVAL 306
DB 766 ACGATGTCATCAGCTTCAGCAGTCAACCAAGTGGCTTCACTTCAGCAGTCAAC 825
QY 307 THRTHRVALATHRGLINGLYVALSERGLYVALTHRPROARGSERGLYVALILE 326
DB 826 GCTTCAGCAGTCAACCAAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCA 885
QY 327 ILETHRVALAVALAVALAVALAVALAVALAVALAVALAVALAVALAVALA 344
DB 886 ACAAGTGGCTTCACTTCAGCAGTCAACCAAGTGGCTTCACTTCAGCAGTCA 939

```

```

RESULT 24
US-08-961-527-370
Sequence 370, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 370:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-370

Alignment Scores:
Pred. No.: 9.41e-08 Length: 750
Score: 165.50 Matches: 62
Percent Similarity: 38.95% Conservative: 49
Best Local Similarity: 21.75% Mismatches: 131
Query Match: 6.99% Indels: 43

```



```

Oy      389 SerLeuThrSerLysTrpValThrValGlyValGlyAlaValAlaAlaProAlaLeu 408
       ::::|||||:::   ::|||                                     |||||
Db     12286 GCGCTTACGAACGGGTATTCTGTCGCCGG------ACACTGGCGTTG 12245
Oy      409 GlyLysGlyIleMetGlnMetGln---LeuSerGluMetGln-----GlnAsn 423
       :|||:::|||||  |||:|||:::|:::|                         |||||
Db     12244 AATATATAGAGCTGGCGCAATCCAAATGTGTGTCTCAGTTACTGCTTTCCGGGTCAAAC 12185
Oy      424 ValAlaGlnPheGlnIleGlyValGlyIle-----LysLeuGlnAlaAlaAla 438
       ::::|:::|:::|:::|:::|:::|:::|                       :|||  |||||
Db     12184 TTAAC TGACAGACAGAGTGAAGAAGCTTAGTGAGAGTGCAGAACGCTCGAGCTGAATAAGCGGCA 12125
Oy      439 AspheIleSerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSer 456
       ::::|:::|:::|:::|:::|:::|:::|                       :|||  |||||
Db     12124 AGCGAGTTACAAAATTCAGCGGAGCTATTTCACAAAGTGTTCACACTGAGATGCC 12071

RESULT 22
US-08-961-527-263
Sequence 263, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 263:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-263

Alignment Scores:
Score:          9e-07           Length:        3744
Percent Similarity: 167.00         Matches:         64
Best Local Similarity: 42.36%      Conservative:    145
Query Match:      22.22%           Mismatches:     145
DB:               7.05%            Indels:         21
                4                  Gaps:          4
US-08-961-527-263 (1-496) x US-08-961-527-263 (1-3744)
Oy      24 ThrSerThrPrroGInglyValPrroGInGlnAsplYsLeuSerGlnAengluThrLysGln 43
       ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     2924 AGCACGACCACATCATCTTACACAGTCTTAAGTCACTCTTCACAGTAAGTCTGTAACAA--- 2980
Oy      44 IleGlnGlnThrArgGlnGlnGlyLysAsnThrGlnUmetGlnSerAspAlaThrIleAlaGly 63

```

```

Db 2981 -----:::|||||::|
Oy 64 AlaSerGIyLysAspLysThrSerSerThrLysThrLysThrAlaAlaProGInGIngly 83
Db 3014 GCTTCAGCATCAACACAGTGCCTCGGCTTCACGCCATCAACACAGTGCCTTCAGCAAGT 3073
Oy 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 3074 ACCAGTGCCTTCAGTTCACATCAACAAGTGCCTTCACAGCTTCACATCAAGTGCCTCG 3133
Oy 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 3134 GCTTCAGCAACACATCAAGCATCGATCGAATCAAGCCGCAACAGTGCCTTCAGCAAGT 3193
Oy 124 GluGluAlaLysSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuAlaAla 143
Db 3194 ACCAGTGCCTTCAGTTCAGTTCAGTCAACACAGCGCCCTCGGCTTCACACACACTGCTTC 3253
Oy 144 GlnMetLysGluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 3254 GAATCGGCTTCACACAGCGCTTCGCGCTTCAGCAACAGCACTGCTTCGATCGGCTCA 3313
Oy 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
Db 3314 ACCAGCGCTTCAGCC-----TCAGCATCAACAGAGGCTTCGCTTCAGCA 3358
Oy 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThrLysSerAla 203
Db 3359 AGCACACAGCGCTTCGGGCTTCAGCATCAACAGATCGTTCAGCTTCAGGCTCAACACAGTCT 3418
Oy 204 LeuSerAsnTYraLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
Db 3419 --TCAGCTTCAGCATCAACAAGTGCCTTCAGCAAGATCAAGATCGCTTCGATCGG 3475
Oy 224 LysGlnAlaIleLysIleAspLysGluArgGluIuArgGlnIuArgGlnMetLysAlaAlaGlu 243
Db 3476 CATCAACAGCGCGCTTC-----AGTAGCATCAACAGATCGTTCAGCTTCAG 3523
Oy 244 GlnLysSerLysAspLeuGlu-GlyThrMetAspThrValAlaSerValMetIleAlaVala 263
Db 3524 CAGAGCACTTCAGCTTCGATCGCTTCACCGCTTCACACAGTGCCTTCACCTCAACACAGAGCC 3583
Oy 263 IserValAlaIleThrValIleSerIleValAlaAlaIleIlePheThrCysGlyAlaGlyLe 283
Db 3584 CTCACCTTCACAGATACCATCGCTTCAGCTTCAGCTTCAGCTTCAGCAAGTGCCTTCAC 3643
Oy 283 uAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaAlaGlyAlaAlaGlyAl 303
Db 3644 CAGTGCATCTGAATCGGATCAACACAGTGCCTTCAGCTTCAGCAAGTCAAGTCAATGCATCGC 3703
Oy 303 aAlaAlaAlaThrThrValAla 310
Db 3704 TTCACATCAACACAGTGCCTCG 3725

RESULT 23
: Sequence 368, Application US/08961527
: Patent No. 6420135
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
: COMPUTER: HP Vectra 486/33

```







Sequence 1, Application US/08591079  
Patent No. 5972899  
GENERAL INFORMATION:  
APPLICANT: Zychlinsky, Arturo  
APPLICANT: Chen, Yajing  
TITLE OF INVENTION: Apoptosis Induced by Shigella ipab  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/591,079  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Livanat, Shmuel  
REGISTRATION NUMBER: 33,949  
REFERENCE/DOCKET NUMBER: 15661-20017.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0764  
TELEX: 90-4030 MRSNFOERSMH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1743 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1743  
NAME/KEY: misc\_feature  
LOCATION: 1..1743  
OTHER INFORMATION: /note="Shigella flexneri ipab"  
US-08-591-079-1  
Alignment Scores:  
Pred. No.: 3,936-09 Length: 1743  
Score: 185.00 Matches: 105  
Percent Similarity: 38.10% Conservative: 84  
Best Local Similarity: 21.17% Mismatches: 171  
Query Match: 7.81% Indels: 136  
Gaps: 21  
US-09-889-314-2 (1-496) x US-08-591-079-1 (1-1743)  
QY 42 LysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMet----- 55  
DB 316 AAGTCCACGACAGGACAGACAGCAAAAAAACCTGAAATTCCTCGATAAATTACACT 375  
QY 55 ----- 55  
DB 376 CTTCTATCTGAACCTGACGACTAACCCAGACTATGAAAAACAATTAATAACTAAA 435  
QY 56 GluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrLys 75  
DB 436 AACGCGATTCTTAATAAAGACCTAGAAATAAATTAAACCAATTCACAAACAGATTA 495  
QY 76 ThrGluThrAlaProGlnGlnGlnValAlaIleAlaGlyLysGluSerSerGluSerGlnLys 95  
DB 496 TCGAACCTCGATCA-----GAGTCACCAAGAAAAAGAAAA 531  
QY 96 AlaGlyAlaAspThrGlyValSerGlyAlaIleAlaThrThrAlaSerAsnThrAlaThr 115

DB 532 TTAGCCGGGAGAA-----ATACAACTCACTATCAAAAAAGCCAGCAGTT 579  
QY 116 LysIleAlaMetGlnThrSerIleGluAlaSerLysSerMetGluSerThrLeuGlu 135  
DB 580 AAA-----GACAGCATGATTGAGCAAGAAACCTGTCATTCATCAGC----- 624  
QY 136 SerLeuGlnSerLeuSerAlaIleGlnMetLysGlnValGluAlaValAlaIleAla 155  
DB 625 AAACCTTACAGATAATCAATGCACTCCAAAAGAAATGACTCT-----TTTCTGCA 678  
QY 156 LeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGly 175  
DB 679 TTTTCAACACAGCATCT--GCTGACAGCTATCAACCCAGCAAGAAATCATTAACCGCA 735  
QY 176 ValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195  
DB 736 CTTGCCAGTGTACTCAATTTG-----ATGGCAACCTTATTCATCACTA 777  
QY 196 LeuGlyGluAlaThrLysSerAlaLeuSerAsnTyraLaserThrGluAlaGlnAlaAsp 215  
DB 778 GTTGGAAAAAATATGAGAAATCTTTAAAAATGATCGCTCTATTCCAGTCTCTCCA 837  
QY 216 GlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluGlu 235  
DB 838 GAATCAGAAAAACGAAATGAGAGAAATCT-----GATGAG 876  
QY 236 TyrGlnGluMetLysAlaAlaGlu--GlnLysSerLysAspLeuGluGlyThrMetAsp 254  
DB 877 TAT-----GCTGCTGAAGTACTTAAACGAGAAAGACTAACACAGACTAATGGGT 924  
QY 255 ThrValAsnThrValMetIleAlaValaLaserValAlaIleThrValIleSerIleValAla 274  
DB 925 TGTGTTGGAAAAATACTTGGGCACCT-----TTACTATCTGTATGTTGTGTA 975  
QY 275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaIleAlaValaGlyAla 294  
DB 976 GCAGCTTTTCTGAGAGACCTCTCTAGCAGCGACGTGTTGTTACTCTCT----- 1029  
QY 295 AlaAlaIleAlaGlyAlaAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleThr 314  
DB 1030 -----ATGCTT 1035  
QY 315 ValGlnAlaValaIleGlnAlaValLysGlnAlaValIleThrAlaValaGlnAlaIle 334  
DB 1036 ACGGATGCTATGATGACACAGCAGCAGCCGC----- 1065  
QY 335 ThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuVal 354  
DB 1066 AATTCCTTATGGAACACAGCCGTAATCGATCATGAAGACAGTCTTGAACCCCTAATC 1125  
QY 355 LysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGln 374  
DB 1126 AAACCTCTTCAGATGCAATTCACAAA-----ATGCTCGAAGGCTTGGGC 1170  
QY 375 MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrp 394  
DB 1171 GTGACTCGAAAAAGCCAAAATGATTGGCTCTATTCTGGGGCAATTCGACGCGCTCTT 1230  
QY 395 ValThrValGlyValGlyValValAlaIleAlaIleAlaIleGlyLysGlyIleMetGln 414  
DB 1231 GTCTAGTTCGACGACGCTGTTCTCGTAGCC-----ACGTGTGTTAA----- 1272  
QY 415 MetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeu 434  
DB 1273 -----CAGGCAGCAGCAAAAACCTGCAAGAAATTTGGCAAAATA 1311  
QY 435 -----GlnAlaAlaIleAspMetIleSerMetPheThrGln--PheThrGlnAla 451  
DB 1312 ATTAGTAAACCTCAGACACCTTATACCAAGATTCTCAAGATTCTTCTCACTG 1371  
QY 452 SerLysIleAlaSerLysGlnThrGlyGluSerAsnGluMet----- 465

APPLICANT: NORRIS, STEVEN J.  
 APPLICANT: JING-REN, ZHANG  
 APPLICANT: HARDHAM, JOHN M.  
 APPLICANT: HOMELL, JERRILYN K.  
 APPLICANT: BARBOUR, ALAN G.  
 APPLICANT: WEINSTOCK, GEORGE M.  
 TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELLIA  
 FILE REFERENCE: US/09/125,619  
 CURRENT APPLICATION NUMBER: US/09/125,619  
 CURRENT FILING DATE: 1999-01-27  
 NUMBER OF SEQ ID NOS: 48  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 7766  
 TYPE: DNA  
 ORGANISM: Borrelia burgdorferi  
 FEATURE:  
 NAME/KEY: modified\_base  
 LOCATION: (127)  
 OTHER INFORMATION: R = A OR G  
 US-09-125-619-3  
 Alignment Scores:  
 Pred. No.: 3,52e-08 Length: 7766  
 Score: 186.00 Matches: 129  
 Percent Similarity: 37.00% Conserved: 76  
 Best Local Similarity: 23.29% Mismatches: 223  
 Query Match: 7.85% Indels: 126  
 Gaps: 25  
 DB:  
 US-09-889-314-2 (1-496) x US-09-125-619-3 (1-7766)  
 OY 5 SerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValLeuThr 24  
 Db 52 GCTGTAGTCTGTTAGTGGG-----GAGCAGATATTAAAGTCGATTGTTAAG 99  
 OY 25 SerThrProGlnIlyValProGlnIlyAspIlyLeuSerGlyValGlnIlyThrIleGln 44  
 Db 100 GCTGCT-----GGTGAAGCTGCCACAGATGAGARAGAGCGTGGAGGCTAAAGATCCG 153  
 OY 45 GlnGlnThrArgGlnIlyValAsnThr-----GluMetGlu 56  
 Db 154 ATTCGTCTGCTATTGGTGAAGGCTAATGAGAGTGTCCGAGTTTAAGATGAGATGAAG 213  
 OY 57 SerAspAlaThrIleAlaGlyAla-----SerGlyIlyAspIlyThrSer 71  
 Db 214 AAGGATGATCAGATTGCTGCTATTCCTTTGAGGGGATGGCTTAAGGAGCTGGGAGTTG 273  
 OY 72 SerThrThrIlyThrGluThrAlaProGlnIlyVal-----AlaAlaGlyIlyGlu 89  
 Db 274 GCGTGAAGATGATGAGCAAGGAGGCTGAGGGGCTTAAAGGAGCTGGGAGTTG 333  
 OY 90 SerSerGluSerGlnIlyAla-----GlyAlaAspThrGlyValSerGly 104  
 Db 334 TTGATTAAGCTGTAAAGCTTAAAGACAGCTGAGGGGCTTCAAGTCTGCTGCA 393  
 OY 105 AlaAlaAlaThrThrAlaSerAsnThrAlaThrIlyIleAlaMetGlnThrSerIleGlu 124  
 Db 394 ATTGGAAGATTGTGCTGATGATTAATGCTGCAAGGTTCTGATTAAGGAGCTGTGAAG 453  
 OY 125 GluAlaSerIlySerMetGluSerThrLeuGluSerIleu-----GlnSerLeuSer 141  
 Db 454 GCGATTCTTAAGGGGATTAAGGAGATTTTGAAGCTGTGGGGGAGTAAAGCTGAAA 513  
 OY 142 AlaAlaGlnMetIlyGlu-----ValGlnAlaValAlaValAla 154  
 Db 514 GTTCTGCTGCTTAAGAGGCAATGAAAGCAAGGAGAGTTGTTTGGGAAAGTTGATGCT 573  
 OY 155 AlaLeuSerGlyIlySerSerGlySerAlaIlyLeuGluThrProGluLeuProIlyPro 174  
 Db 574 GCTATGCTGGGACAGGAGGCTGCTGTCAGAG-----606  
 OY 175 GlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaIlyAlaIleGln 194

Db 607 -----GCGGCTGGCTGTAGT 624  
 OY 195 ThrLeu-----GlyGluAlaThrIlySerAlaLeuSerAsnIlyAlaSerThrGlnAlaGln 213  
 Db 625 GCTGTAGTGGGAGCAGATATTAAAGCTGTTAAAGCTGGTGGCGCTGCTGCT 684  
 OY 214 AlaAspGlnThrAsnIlyLeuGluIlyGln-----AlaIleIlyIleAspIly 231  
 Db 685 GATCAGAGGAGGAAAGAACCTGGGATGCTAAATCCGATTCGCTGCTGATTTGGAG 744  
 OY 232 ---GluArgGlnIlyThrGlnIlyMetIlyAlaAlaGlnIlySerIlyAspLeuIly 250  
 Db 745 GCTATGCGGAGATGTGTGGAGATTAAATCATGATGAGGAGAAAGAGATGAT 798  
 OY 251 GlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIle 270  
 Db 799 -----CAGATTGCTGCTCTATTTCCTTTGAGCGGGATG 831  
 OY 271 SerIleValAlaAlaIle---PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIly 289  
 Db 832 GCTAAGATGCAAGTTCGCTGTGAAGTGTGTGTGAGAGGAGGAGCTGAGGG 891  
 OY 290 AlaAlaValGlyAlaAla-----AlaAlaGlyGlyAlaAla 301  
 Db 892 GCTATTAAAGGAGCTGCTGAGTGTGTGATTAAGCTGTAAGACAGCTGAG 951  
 OY 302 GlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAla---ValValGln 320  
 Db 952 GGGCTTCAAGTGTGATGATCATTTGAGAGATTTGGCTAATGCTGCTGCTCAAG 1011  
 OY 321 AlaValIlyGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaIleIlyAla 340  
 Db 1012 GTTGTGATTAAGGAGCTGAGTGTGAGGAGTGTGTAAGGAGATTAAGAGATTTGAGGCT 1071  
 OY 341 AlaValIlySerGlyIlyLeuAlaPheIlyThrLeuValIlyAlaIleAlaIlyAla 360  
 Db 1072 GCTGGGGGAGTAA-----AAGCTGAAGTGTGCTGCTACAGAGGGAG 1116  
 OY 361 IleSerIlyGlyIleSerIlyValPheAlaIly---GlyThrGlnMetIleAlaIlyAsn 379  
 Db 1117 AGTATTAAGGGGAGGAGGAGATTTTGGAGAGCTGCTGCTGCTAATTCCTGGGAC 1176  
 OY 380 PheProIlyLeuSerIly-----ValIleSerSerLeuThrSerIlyThrValThr 396  
 Db 1177 AGTGAAGCTGTACCAAGCGGCTGCTGCTTAACTGCTGTTAGTGGAGAGATATTA 1236  
 OY 397 ValGlyValGlyValAlaValAla-----404  
 Db 1237 AGTGGATTT---GTTAAGGCTGCTGATCGGCTGATCAGAGGAGAAAGAGCTGGGAT 1293  
 OY 405 -----AlaProAlaLeuGlyIlyGlyIlyIleMetGlnMetGlnLeuSerGlu 419  
 Db 1294 GCTACAAATCCGATTGCTGCTGCTATTGGAGGCT-----1329  
 OY 420 MetGlnGlnAsnValAlaGlnPheGlnIlyGluValGlyIlyLeuGlnAlaAlaAlaAsp 439  
 Db 1330 AATGAGCAAGATGTGCGGAGTTTAAGATGATGAGAGAGATGATCAGATTGCTGCT 1389  
 OY 440 MetIleSerMetPheThrGlnPheThrGlnAlaSerIlyIleAlaSerIlyGlnThr 459  
 Db 1390 GCTATTGCTTTG---AGGGGATGCTGATAGATGAGTGAAGATTGCTGCAAGATGCT 1443  
 OY 460 GlyIlySerAsnGluMetThrGlnIlyAlaThrIlyLeuGlyAlaGlnIleLeu-----477  
 Db 1444 GGTGAGAAAGGAGAG---GCTGAGGCGCTATTAAAGGAGCTGCTGATTTGTTGATTAAG 1500  
 OY 478 -----LysAlaThrAlaAlaIleSerGlyAlaIleAlaGly 489  
 Db 1501 CTGGTAAAGCTGTAAAGACAGCTGAGGGGCTTCAAGGCT 1542  
 RESULT 18  
 US-08-591-079-1

Oy 447 eTrpInglInlaSerIysIleAlaSerIysInThrIyGluSerAnGluMetThrG1 467  
Db 1634 ----GCATCAACAAGTGGCTCGGCTTCAGCAAGCACATCATCGATCGATCAAC 1689  
Oy 467 nLysAlaThrIySLeuGlyAlaGlnIleuLysAlaTyrAlaIleSerGlyAla11 487  
Db 1690 CAGTGGCTTCAGCAAGTACAGTGTCTCAGCTTCAGCATCAACGCGCTCGGC 1749  
Oy 487 eAlaGlyAla 490  
Db 1750 CTCAGCAAGC 1759  
RESULT 16  
US-08-961-527-365  
; Sequence 365, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8512  
; TELEFAX: (301) 309-8504  
; INFORMATION FOR SEQ ID NO: 365:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1436 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-527-365  
Alignment Scores:  
Pred. No.: 1.43e-09 Length: 1436  
Score: 188.00 Matches: 79  
Percent Similarity: 37.30% Conservative: 59  
Best Local Similarity: 21.35% Mismatches: 214  
Query Match: 7.94% Indels: 18  
DB: 4 Gaps: 6  
US-09-889-314-2 (1-496) x US-08-961-527-365 (1-1436)  
Oy 2 ThrAsnMetSerIleSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGln 21  
Db 298 ACAAGTGGCTTCAGCAAGTACCTCAGCTGGAATCAGCAACGAGATGATCA 357  
Oy 22 ValLeuThrSerThrProGlnGlyValProGlnGlnAspIlySerGlyAsnGluThr 41  
Db 358 GCTTCAGCATCAACA-----AGTGGCTTCAGCTTCAGCAAGTATCTCAGCGCTGATCG 411

Oy 42 LysGlnIleGlnInThrArgInGlyLysAsnThrGluMetGluSerAspAlaThrIle 61  
Db 412 GCATCAACGAGGCGCTCGGCTTCAGCAAGTACTACGGCTCAGCATCAACGCAACAGT 471  
Oy 62 AlaGlyAlaSerGlyLysAspIlyThrSerThrThrIySThrGluThrAlaProGln 81  
Db 472 GCTTCGGCTTCAGCAAGTACAGTGTGAGTACAGTCAACAGTACAGTACAGTAC 531  
Oy 82 GlnGlyValAlaIleAlaGlyLysGluSerSerGluSerGlnIlyAlaGlyAlaAspThrGly 101  
Db 532 GCAAGCAGATCAGCTTCGAATCTGATTCATCACCAGTGCCTCAGCTCAGCATCAACG 591  
Oy 102 ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla----- 114  
Db 592 GCCTCAGCTTCAGCAAGTACAGTGTGAGTACAGTGTGAGTACAGTGTGAGTAC 651  
Oy 115 -----ThrLysIleAlaMetGlnThrSerIle---GluGlnAlaSerIySerMetGlu 131  
Db 652 GCGTCGACAGTGGCGCTTCACACAGTGCATCATCGCAATCGCAACACAGTGGCTCA 711  
Oy 132 SerThrLeuGluSerIleuSerIleuSerAlaIleGlnIleMetLysGluValGluAlaVal 151  
Db 712 GCCTCAGCAAGTACTAGCGCTTCAGCTCAGCATCAACGAGTGCCTCGCTTCAGCAAGT 771  
Oy 152 ValValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeu 171  
Db 772 ACTAGTCATCAGCTTCAGCAAGTACTACGCTCAGCTCAGCTCAGCTCAGCTCAGCTCA 831  
Oy 172 ProLysProGlyValThrProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLys 191  
Db 832 GCTTCAGCAAGTACAGTGCCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCA 885  
Oy 192 AlaIleGlnThrLeuGlnIlyGlnAlaThrIySLeuAlaLeuSerAsnThrIySerThrGln 211  
Db 886 GCAAGTACTCAGCGCTTCAGTATCAGCATCAACAGTGGCTTCAGCATCAACAGT 945  
Oy 212 AlaGlnAlaAspGlnThrAsnIySLeuGlyLeuGlu-----LysGlnAlaIleLys 228  
Db 946 GCATCAGCTTCAGCAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTCA 1005  
Oy 229 IleAspLysGluArgGluGluGluGlnIyGlnIleMetLysAlaAlaGluGlnIlySerIyAsp 248  
Db 1006 GCATCAACGAGTGCCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAACAGT 1065  
Oy 249 LeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThr 268  
Db 1066 GCCTCGGCTTCAGCAAGTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTCA 1125  
Oy 269 ValIleSerIleValAlaIleAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAla 288  
Db 1126 GCATCAACGAGTGCCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAACAGT 1185  
Oy 289 GlyAlaAlaValAlaIleAlaIleAlaIleGlyAlaIleGlyAlaAlaAlaIleAlaIleThr 308  
Db 1186 GCGTCAGCTTCAGCAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTCA 1245  
Oy 309 ValAlaIleThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaIleThr 328  
Db 1246 GCATCAACGAGTGCCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAACAGT 1305  
Oy 329 AlaValArgGlnAlaIleThrAlaIleAlaIleLysAlaAlaValLysSerGlyIleLysAla 348  
Db 1306 GCGTCGGCTTCAGCAAGTACAGTGTGCG---TCAGCTTCAGCAACGACAGTGTGTCAGCT 1362  
Oy 349 PheIleLysThrLeuValLysAlaIleAla 358  
Db 1363 TCAGCATCAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTCA 1392  
RESULT 17  
US-09-125-619-3  
; Sequence 3, Application US/09125619  
; Patent No. 6437116  
; GENERAL INFORMATION:

```

US-08-961-527-364
; Sequence 364, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae polynucleotides and sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 364:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2550 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-961-527-364
;
; Alignment scores:
; Pred. No.: 1.43e-09 Length: 2550
; Score: 192.00 Matches: 110
; Percent Similarity: 33.97% Conservative: 68
; Best Local Similarity: 20.99% Mismatches: 275
; Query Match: 8.10% Indels: 71
; Gaps: 10
;
US-09-889-314-2 (1-496) x US-08-961-527-364 (1-2550)
OY 5 SerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleThr 24
Db 289 TCACACAGTGGCTCGCTCGACATCGACAGCCCTCAGCTTCAGCAAGTACCACTGCT 348
OY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIle 44
Db 349 TCAGCCTCAGCCTGACAGAGTGGCTCGGCTTCACACAGTGCATGATCGCATCAACC 408
OY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
Db 409 AGTGGCTCAGCCTCAGCAGTACTAGTCATCAGCTTCACATCAACGAGTGCATCGGCT 468
OY 65 SerGlyLysAspLysThrSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
Db 469 TCAGCATCAACAGTGGCTCGGCTTCAGCGTCAACAGTGGCTCAGCTTCAGCAAGTACC 528
OY 85 AlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSerGly 104
Db 529 AGTGCTTCAGTTCACAGCATCAACAGTGGCTCGCTCAGCATCGCAAGTGCCTCGGCT 588
OY 105 AlaAlaIleThrThrAlaSerAsnThrAla---ThrLysIleAlaIleMetGlnThrSerIle 123

```

```

Db 589 TCAGCAAGCATCATCAGATCTGAATCAGGCTCGACAGAGTGGCTCGGCTCAACAGTACA 648
OY 124 ---GlnGlnLaserLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerIle 142
Db 649 TCCTAATCGGCATCAACCACTGGCTCGCTCAGCAAGTACTGATCGATCACTTCACGA 708
OY 143 AlaGlnMetLysGlnValGlnAlaValAlaValAlaIleAsnSerGlyLysSerGly 162
Db 709 TCACGAGTGCATCGGCTTCGGCTCAGCAAGTGCATCGATGACAGTCAAGTCAAGTGGC 768
OY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182
Db 769 TCAGCTTCGGCATCAACAGT----- 789
OY 183 IleGlnIleGlyLeuAlaIleAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSer 202
Db 790 -----GCTCGCTTCAGCAAGTCAACGATCAGCATCTGATATGAGGCTCAAC 834
OY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
Db 835 AGTGGCTTCGGCTTCAGCAAGTACAGTGCCTTCAGCTTCAGCATCAACC-----AGCGCC 888
OY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlyThrGlnLysMetLysAla 242
Db 889 TCGGCTCAGCAAGCAGCTCAGCTTCGATGCGGCTCAACAGCGCTCGGCTCGACGA 948
OY 243 GlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
Db 949 AGCACTCAGCTTCGATCGGCTCAACGAGCTCAGCTCAGCATCAACAGTGC 1008
OY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyValGly 282
Db 1009 TCGGCTTCAGCAAGCAGCGCTCGGTCAGTCAACAGATGATCTCAGCTTCACGC 1068
OY 283 LeuAlaGlyLeuAlaIleAlaIleAlaValAlaGlyAlaAlaIle-----Ala 297
Db 1069 TCACACAGTGCCTTCAGCTCAGCATCAACAGGCTCAGCTCAGCAAGTATCTCACGC 1128
OY 298 GlyLysAlaAlaGlyAlaAlaAlaIleThrValIleThrGlnIleThrValGlnAla 317
Db 1129 TCTGATCGGATCAACGATCGGCTCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1188
OY 318 ValValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla 337
Db 1189 AGCACAAGTGTTCAGCTCAGCAAGTATCTCAGCGCTGATATGCGCATCAACAGTGGC 1248
OY 338 IleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
Db 1249 ---TCCGCTTCAGCAAGTACTAGCGCTCAGCAAGTCAAGTCAAGTCAAGTCAAGTCA 1305
OY 358 AlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377
Db 1306 CGC-----TCACAGATGCTGTCTGATCAGTCAACAGTCAAGTCAAGTCAAGTCAAG 1356
OY 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397
Db 1357 AGCAATCATGCTTCGATTCGATTCATCAACAGTGGCTGATCGCTCAGATCGACAGCGCC 1416
OY 398 GlyValGlyValValAlaIleAlaIleProAlaLeuGlyLys-GlyIleMetGlnMetGln 417
Db 1417 TCAGCTTCAGCAAGTACAGTGCCTGATCGCTCAGCAAGTATACCAATGAGTGGCTCGCGC 1476
OY 417 userGlnMetGln-----GlnAsnValAlaGlnPhe 427
Db 1477 TCACAAAGTGGCTCGGCTCAACAGTGCATCTGAATCGGATCAACAGTGGCTCGAGCT 1536
OY 427 eGlnLys----- 429
Db 1537 CAGCAAGTACTAGTGCATGAGCTTCAGATCAACAGATGATCGGCTTCGGGCTCAACCA 1596
OY 430 ---GlnValGlyLysLeuGlnAlaIleAlaIleAspMetIleSerMetPheThrGlnPhe 447
Db 1597 GTGCATGAGATGACAGCAAGTACCAAGTGCATCTCC----- 1633

```



NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESS: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 367:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1702 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-367  
Alignment Scores:  
Pred. No.: 5,87e-10 Length: 1702  
Score: 193.00 Matches: 112  
Percent Similarity: 37.85% Conservative: 75  
Best Local Similarity: 22.67% Mismatches: 267  
Query Match: 8.15% Indels: 42  
Gaps: 9  
US-09-889-314-2 (1-496) x US-08-961-527-367 (1-1702)  
QY 19 MetserglValIleuthrSerThrProglInglyAlaProglInglAspLysLeuSerGly 38  
Db 279 CTTCAGCAAGTACAGTGCCTCAGCTCAGCAAGTACAGTGCCTCAG----- 326  
QY 39 AsnGluThrLysGlnIleGlnInThrArgGlnGlyLysAsnThrGluMetLysAsp 58  
Db 327 -----CCTCAGCGTCGACCAAGTGCCTCAGCC-TCACCACTGCATCTGAATCGGCA 376  
QY 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThr 78  
Db 377 TCACCAAGTGCCTCAGCTCAGCAAGTACAGTGCATCAGCTCAGCAAGTACAGTGC 436  
QY 79 AlaProglInglLysAlaAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAla 98  
Db 437 TCGGCTTCAGCATCACCAGTGCATCAGTGCATCAGTGCATCAGTGCCTCAGTTCGCA 496  
QY 99 AspThrGlyAlaSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla----- 114  
Db 497 TCACAAGTGCCTCAGCTCAGCAAGTACAGTGCCTCAGCTCAGCTCAGCAAGTGCCT 556  
QY 115 -----ThrLysIleAlaMetGlnThrSerIle-----GluGluAla 126  
Db 557 TCAGCTTCGCGTCACAGGCGCTCGCGCTCAGCAAGTATCTCAGCGCTCGATTCGGCA 616  
QY 127 SerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLys 146  
Db 617 TCACAAGTGCCTCAGCTCAGCAAGTACAGTGCATCAGTGCATCAGTGCATCAGTGC 676

QY 147 GluValGluAlaValAlaAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeu 166  
Db 677 TCGGCTCAGCAAGTACAGGCGCTCAGTGCATCAGTGCATCAGTGCATCAGTGCAT 726  
QY 167 GluThrProglInglLeuProLysProglYValThrProArgSerGluValIleGluIleGly 186  
Db 737 AGTACCTCAGCATCTGATCAGCATCAGCAAGTGCATCAGTGCATCAGTGCATCAG 790  
QY 187 LeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSerAsn 206  
Db 791 AGTCTTCAGCTCAGCAAGTACAGTGCATCAGTGCATCAGTGCATCAGTGCATCAG 850  
QY 207 TyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLys-LeuGlyLeuGlnLysGlnAl 226  
Db 851 TCAGCAAGTACAGTGCCTCAGCATCAGCTCAGCAAGTGCCTCAGTGCATCAGTGCAT 910  
QY 226 AlieLysIleAspLysGlu-----ArgGluGluTyrGlnGluMetLysAlaAlaG 243  
Db 911 AGTCCGTGAGTCAGCATCAGCATCAGTGCATCAGTGCATCAGTGCATCAGTGCAT 969  
QY 243 uGlnLysSerLysAspLeuGlnLysThrMetLysPheThrValAsnThrValMetIleAla 263  
Db 970 ATCTGCATCAGCAAGTGCCTCAGCATCAGTGCATCAGTGCATCAGTGCATCAGTGC 1029  
QY 263 LserValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLe 283  
Db 1030 CAGTGCCTCA-----GCCCTCAGCAAGTACAGTGCCTCAGTGCCTCAGTGCCT 1068  
QY 283 uAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAl 303  
Db 1069 AGCGTCAGCAAGTGCCTCAGCATCAGTGCATCAGTGCATCAGTGCATCAGTGCATC 1128  
QY 303 aAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 323  
Db 1129 AGCTCAGCAAGTACAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTC 1179  
QY 323 sGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 343  
Db 1180 TTCAGCAAGTACAGTGCCTCAGCATCAGTGCATCAGTGCATCAGTGCATCAGTGC 1226  
QY 343 sSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 363  
Db 1237 TACAGCGCTCAGCTCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAG 1296  
QY 363 sGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLe 383  
Db 1297 AGCTCAGCAAGTACAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTC 1347  
QY 383 uSerLysValIleSerSerLeuThrSerLysThrValThrValGlyAlaGlyAlaVal 403  
Db 1348 CTCAGCTCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGT 1407  
QY 403 lAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 423  
Db 1408 GAGTGCATCAGCTCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGC 1467  
QY 423 nValAlaGlnPheGlnLysGluValAlaGlyLysLeuGlnAlaAlaAlaPheIleSer 443  
Db 1468 TTCAGCATCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTC 1527  
QY 443 tPheThrGlnPheThrProglInglAlaSerLysIleAlaSerLysGlnThrGlyGlnSer 463  
Db 1528 AAGTGCCTCAGCTCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGC 1587  
QY 463 nGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaLysAlaAla 483  
Db 1588 TTCAGCATCAGCAAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTCAGTGCCTC 1647  
QY 483 eSerGlyAlaIleAlaGly--AlaLysLysThrAsnAsn 495  
Db 1648 GAGTGCCTCAGCTCAGCAAGTGCATCAGTGCATCAGTGCATCAGTGCATCAGTGC 1685  
RESULT 15

COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/961,527  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, A. Anders  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PB340P1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 363:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4483 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-961-527-363

Alignment Scores:  
 Pred. No.: 6,98e-10 Length: 4483  
 Score: 199.00 Matches: 111  
 Percent Similarity: 37.80% Conservative: 78  
 Best Local Similarity: 22.20% Mismatches: 274  
 Query Match: 8.40% Indels: 39  
 Gaps: 11

US-09-889-314-2 (1-496) x US-08-961-527-363 (1-4483)

QY 2 ThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21  
 DB 689 ACAAGTGGCTGAGCTTCAGCAAGTACTGTGCATCAGCTTCAGCAAGTGCATCG 748  
 QY 22 ValIeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThr 41  
 DB 749 GCTTGCGGCTCAAC-----AGTGCATCAGATCA 778  
 QY 42 LysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIle 61  
 DB 779 GCAAGTACCAAGTGGCTTCAGCAAGTGCCTGCGCTTCAGCAAGCAGT 838  
 QY 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGln 81  
 DB 839 GCGTGGCTTCAGCAAGTACTAGCGCTCAGCTCAGCTCAGCTCAGCAAGTGCATCTCA 898  
 QY 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGly 101  
 DB 899 GCAAGTATCTCAGCTGCTGATCGCATCAGCAGTGCCTGCGCTTCAGCAAGTACTAGC 958  
 QY 102 ValSerGlyAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaIleMetGlnThr 121  
 DB 959 GCGTCAACCTTCAGCAAGTCAAGTGCATGCGCTTCAGCTCAACGATCGCT----- 1012  
 QY 122 SerIleGluLysLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSer 141  
 DB 1013 -----GATTCGGCATCAACGAGTGGCTCGCTTCAGCAAGTACTAGCGCTCAGCTCA 1066  
 QY 142 AlaAlaGlnMetLysGluValAlaGluValAlaValAlaAlaLeuSerGlyLysSerSer 161  
 DB 1067 GCGTCAACAGTGCATGCGCTTCAGCAAGTCAACGATGCTCGCTTCAGCAAGTACTAGC 1126  
 QY 162 GlysAlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGlu 181

DB 1127 GCCTAGCCTCAGCTCAACA---AGTGCATGGCTTCAGCTCAGCAAGTGCCTGAG 1183  
 QY 182 ValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrIleuGlyGluAlaThrLys 201  
 DB 1184 TCAGCATCAACAGAGT---GCGTCAGCTTCAGCAAGCAGCATCATCAGCTTCGATCTGCATCA 1240  
 QY 202 SerAlaLeuSerAsnThrLysThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGly 221  
 DB 1241 ACCAGTGGCTCAGCTCAGCATTCAGCAAGCGCTTCAGCTTCAGCAAGCAGTGCCTCA 1300  
 QY 222 LeuGlnLysGlnAlaIleLysIleAspLysGlu-----ArgGlnLysGlnGlu 238  
 DB 1301 GCTCAGGCTGCAAGTGCCTGCTTCAGCAAGTCAAGTGCCTCAGCTCAGCAAGT- 1359  
 QY 239 MetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMetLysPThrValAsnThr 258  
 DB 1360 ACCAGTGGCTCAGCTCAGCTCAGCAAGTGCCTGCGCTTCAGCAAGTGCATTCAGT 1419  
 QY 259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThr 278  
 DB 1420 GCATCAACCAAGTGGCTCAGCTCAGCAAGTACT---AGCGCTCAGCTCAGCAAGTCAAG 1476  
 QY 279 CysGlyValGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaValAlaAlaAlaGly 298  
 DB 1477 AGTGGCTGGCTTCAGCAAGTACTAGTGCATCAGCTTCAGCAAGTACTAGCGCTCAGCC 1536  
 QY 299 GlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaVal 318  
 DB 1537 TCAGCTGCAACAGCGCTCAGT---TCAGCAAGTACAGTGCCTCAGCTCAGCTCAG 1593  
 QY 319 ValGlnAlaValLysGlnAlaValIleThrAlaValAlaGlnAlaIleThrAlaAlaIle 338  
 DB 1594 ACAAGTGGCTGGCTTCAGCAAGTACTCAGCTCAGCTCAGTCAATCAGCATCAACATGGG 1650  
 QY 339 LysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrValLysAlaIleAla 358  
 DB 1651 TCGGCTTCAGCATCAACAGTGCATCAGCTTCAGCAAGTCAACAGTGCCTCAGCTCAGCA 1710  
 QY 359 Lys---AlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377  
 DB 1711 ACTACAGTGGCTGGCTTCAGCAAGTCAACAGTGCCTCAGTCAAGCTCAGCAAGTGCCT 1770  
 QY 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397  
 DB 1771 TCTGAATCCGATCAACAGTGCCTGCGCTTCAGCAAGCAGCAGTCTCGCTTCAGCG 1830  
 QY 398 GlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGluLeu 417  
 DB 1831 TCACAGAGTGGCTGCTGATCAGCATCAACAGTGCCTCAGCTCAGCAAGCAGTCAAGCTT 1890  
 QY 418 SerGluMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLys-LeuGlnAlaAl 437  
 DB 1891 CTGAATTCGATCAACAGTGCCTGCTCAGCTCCG-----ATCAACAGCGCTCGCGCTTC 1944  
 QY 437 AlaAspMetIleSerMetPheThrGlnPheThrGlnAlaSerLysIleAlaSerLys 457  
 DB 1945 AGCA-----AGTCAACAGTGCCTCAGCTCAGC 1971  
 QY 457 GlnThrGlyGluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeu 477  
 DB 1972 ATCAACCAAGTCAAGTGCCTCAGCTCAGCAAGTGCCTCAGCTCAGCTCAGCTCAGTGC 2031  
 QY 477 uLysAlaGlyAlaAlaIleSerGlyAlaIleAlaGlyAlaHisLysThrAsnAspPhe 496  
 DB 2032 CTCGGCTTCAGCAAGTCAACAGTGCCTCAGCTCAGCAAG-CACAGTGGCTCAGCTTC 2088

RESULT 14  
 US-08-961-527-367

; Sequence 367, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ. ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 46819  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 72:  
US-09-453-702B-72

## Alignment Scores:

Pred. No.:	1,23e-08	Length:	46819
Score:	203.00	Matches:	127
Percent Similarity:	37.33%	Conservative:	66
Best Local Similarity:	24.56%	Mismatches:	208
Query Match:	8.57%	Indels:	116
DB:	4	Gaps:	20

US-09-889-314-2 (1-496) x US-09-453-702B-72 (1-46819)

OY 5 SerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnIleValLeuThr 24  
DB 32174 TCAGGCTCTTCACGGCAGCA-----ACGGCA 32200  
OY 25 SerThrProGlnGlyValProGlnIlyAsnGlyValThrIleVal 44  
DB 32201 TCACAAAGGCTACTGAAAGCATCAAAAGTGCTGCCGTCGACAGTCTCCCAAAAGCGCG 32260  
OY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleVal 64  
DB 32261 GCGGCTACCAAGCGCGGTCGCGGCAAAACGTCAGAAACGCAATGCGCAGTGTCCACACA 32320  
OY 65 SerGlyLysAspLysThrSer---SerThrLysThrGlnThrAlaProGln-----81  
DB 32321 TCAGCGCGCATCTGCAATCCACGCGACGCAAGCGTCAGAAAGCTGCTCCTCAGCC 32380  
OY 82 GlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnIlyAlaGlyAlaAspThrGly 101  
DB 32381 AAGGATGGCTGCGCTTCAAAAGAGCGCGCAAAATCATCAGAAACGACGACCTCCGAC 32440  
OY 102 ValSerGlyAlaAla-----AlaThrThrAlaSerAsnThrAlaThrLysIleAlaMet 119  
DB 32441 GCCAGTAGCGAGCTCTCCGCAACGCGGCGCAAGCAATTCGCGCAAG-----GCGGCC 32494  
OY 120 GlnThrSerIleGlnIlyLeuAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSer 139  
DB 32495 AAAAGCTCT---GAGACAAACGCTAAGTCTCTGAAACGCGACGAGAA-----CAGAGT 32545  
OY 140 LeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSerGlyLys 159  
DB 32546 GCGCTCCGCGAGCAGAGGCTCAAAACAGCGGCTGATATCTCCAGCTCCCGCTACACA 32605  
OY 160 SerSerGlySerAlaLysLeuGlnThrProGlnIlyLeuProLysProGlyValThrProArg 179  
DB 32606 AGTGGCGGCGAGCC-----TCAGCCAGTGGCACCGCGCGCGGAAATTCGCGAGAA 32656  
OY 180 SerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAla 199  
DB 32657 AGTGGCGCATCTGCTCTTCAACAGCACAACGAGGCT-----GGCGAGCC 32704  
OY 200 ThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsnLys 219  
DB 32705 ACTGAACAGCGCAGCGACGACGAGTCTGCTCCGCA-----32743

OY 220 LeuGlyLeuGlnIlyGlnAlaIleLysIleAspLysGlnArgGlnIlyThrGlnIlyMet 239  
DB 32744 -----GCG 32746  
OY 240 LysAlaAlaGlnIlyGlnLysSerLysAspLeuGlnIlyThrMetAspThrValAsnThrVal 259  
DB 32747 AAGACATCCGAACGAAACGAAACCGTTGGAAACACGCGCAAGATCTCTCAAAAACG---32803  
OY 260 MetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCys 279  
DB 32804 -----GTCGCGCATCTGCTACCCAGT 32824  
OY 280 GlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValGlnAlaAlaAlaGlyGly 299  
DB 32825 TCGCGCGGCTCATCGCATCTCTGCTCTTCAAAAGATGAGCGCAGCACGACGACG 32884  
OY 300 AlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaVal 319  
DB 32885 TCAGCAGCGAAGACGACGCGCACGCGCATTCACGAGGCGACGAGCGAGCT-----32938  
OY 320 GlnAlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLys 339  
DB 32939 -----GGTAGTCGACGCGCAGCAGCTCAGACGCAAAAGTACGCGGGAATCT 32983  
OY 340 AlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLys 359  
DB 32984 GCAGCA-----ACGCGCGCTGACACACGCGCAAA 33013  
OY 360 AlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsn 379  
DB 33014 CCGGCAAGATATTCATTCGCGCGCTGAGATGCGACGACGACGAAAGAGGG 33073  
OY 380 PheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyVal 399  
DB 33074 ATAGTACAGCTCAGCAGTGCAGCTACAGC---ACTTCCGATCTGCGCGCAAGCCCA 33130  
OY 400 GlyValAlaValAlaAlaAlaProAlaLeuGly-----409  
DB 33131 AAAGCGTTAAGCGCGCTGATGAGCTAAGGGAATACCCGACAGATGACAG 33190  
OY 410 -----LysGlyIleMetGlnMet-----GlnLeuSerGlnMetGln 421  
DB 33191 ACAGCAGAAAGGATGATGCTAGCAACGCGACCAACGACGACATCTGAAATG---33247  
OY 422 GlnAsnValAlaGlnPheGlnIlyGlnValGlyLysLeuGlnAlaAlaAspMetIle 441  
DB 33248 -----CTGGCGGCAACGCCCAAGTCGGA-----AAGCAGCCTATGACCTTGCT 33292  
OY 442 SerMetPheThrGlnPheThrGlnIlyAlaSerLysIleAlaSerLysGlnThrGlyLys 461  
DB 33293 AAC-----GGAAATATATCTCTCAGAGCGTACGACAGCACAAGAAATGTCAG 33346  
OY 462 SerAsnGlnMetThrGlnLysAlaThrLysLeuGlnIlyAlaGlnIleLeuLysAlaThrAla 481  
DB 33347 CTCAGTAGTACCAACAGCAGCATCTGAAACGCTTGCGCGACACCCAAAGCACTGTAA 33406  
OY 482 Ala-----IleSerGlyAlaIleAlaGlyAlaHisLysThrAsn 494  
DB 33407 CGAGCTAATGATATATGCAATGCGGATCTTCTGCTTAAAGTGAAT 33457

RESULT 13  
US-08-961-527-363  
Sequence 363: Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland



Percent Similarity:	41.02%	Conservative:	85
Best Local Similarity:	23.67%	Mismatches:	182
Query Match:	9.22%	Indels:	107
DB:	2	Gaps:	20

Oy	28	glnlglvalproglnglnasplrsleuserglyasngluthrlysglnlileglnlthr	47
Db	908	cagccgatgatgagtcacacaaaagagatgggagcttcagctttccaaaagaattccacagc	967
Oy	48	arglngllylsasnthrglmetgluserasplathrillealaglylasercllyps	67
Db	968	gcctctggga-----gagcctcagagagcgacgagatctctatgaagccgattcaaa	1018
Oy	68	asplystrhserserthrthrlystrhgltuthralaiproglngllyalalaalely	87
Db	1019	-----aaagcggataccgccacagctgttatgacccctgcgcc	1057
Oy	88	lysglusersergluserglnllysala-----glylasapthrglyal	102
Db	1058	aaaaatgcacgcgagccgcaaaattaaattgcaatgcctgcagccggctgacccggctat	1117
Oy	103	sercllyalalaalathrthrallaserasnthralathrlysllealawetlthrser	122
Db	1118	gcacmacctgaagc-----gcg	1135
Oy	123	ileglnglualaserlyssermetgluserthrleuglserleugluserleuserla	142
Db	1136	gtacacacagcccgcaaaaagacgacacagcagcgaagagcccttgatrtagccacgcat	1195
Oy	143	alaglnmetlys-----gluvalglualavalavalalaleuser	157
Db	1196	gcgcagcgtttaaagcagcagcacagccacaaagcgaagccgataacattctcg	1255
Oy	158	glyluserserglyseralalytleugluthrprogluleuprolysproglvalthr	177
Db	1256	accaaaattccaggaagagcgttaattgccctctcacatcagatttccacaggtacagc	1315
Oy	178	proarg---sergluvalilegluileglyleuallaleuallalelysalleglthrleu	196
Db	1316	gatattcgttcacaaattgtcgcccccctacattatcctcatgattttatgagattgng	1379
Oy	197	glyglualathrlysserallaleuserasrtyralaserthrghlaleghlaleuspeln	216
Db	1376	ggcaaaaattacggaagaaacccgcacaaaagatcttgccctt-----	1417
Oy	217	thrsanlytleuglyleugllysglnallelysilleasplysglunarglgluthyr	236
Db	1418	ttcaacagccttgcaagaaagggcctcagagcc---gaaatgaaaaaagaaatccgctaattc	1474
Oy	237	glnleuwetlysalalaaglunlysserlyaspleugllythrmetspthrval	256
Db	1475	cagaaa-----gagacgcgcacaaagccgagaaaacacacccgattttggatgattac	1525
Oy	257	asnthrvalmetillealavalservalalillethrvalilleserillevalalalalie	276
Db	1526	gggaaatgcttcctgcgcgcctg-----ctaacattgtcaccgcttggccctctgt	1576
Oy	277	phetrlycsglyalaglyleuallaglyleuallalealaglyalalavalaglyalalaa	296
Db	1577	ttttaccgctggcgagctatgcgcctgcgcctgcctggagacttgcgcta-----	1624
Oy	297	alaglyglyalaaaglyalalalalalathrthrvalalathrcnlnlethrvalghn	316
Db	1625	-----atgtggcccgataaaatt-----	1642
Oy	317	alavalalaglnalavallysglnalalaval---ilethrvalavalarglnalaliethr	335
Db	1643	-----gtgaagggcgagcagcgagtgctgtttatttcacagcagcgctaaac	1687
Oy	336	alaalattelysalalavalalysserglyllelysalalathrillelystrhrleuvallys	355

```

Db      1688 CCGATTATGACGACTGTG-----CTGAAGCCGTTAATGGAG 1723
OY      356 AAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
       ||| |||||:::|||| ::|
Db      1724 CTGATTGGCAGCGCATTTACCAGCAAGCCGTG-----GAGATTATGAGCGTTC 1768
OY      376 IleAlaLysAsnPhrProLysLeuSerLysValIleSerSerLeuThrSerLysTrpVal 395
       ||| ::| :|::| :|::| :|::| :|::|
Db      1769 GATAAGAAAACGCCAGAGATGCGCGGCAGCATGTTGTGTCGATTGTGCGCTATTGGC 1828
OY      396 ThrValGlyValGlyValAlaVal-----AlaLaPheoAlaLeuGly 409
       ||| ||| |||||:::|||| |
Db      1829 ATGTGTCGGCGTCATTGTGTGTGTGTCGACTTGTCGGCAAAGCCCGCGCGCAAACTGGGT 1888
OY      410 LysGlyIleMetGlnMetGlnLeuSerGluMetGlnInAsnValAlaGlnPheGlnLys 429
       ::| ::| ||| ::| :|::| :|::| :|::|
Db      1889 AACGCCGCTGAGCAAAATGT-----ATGGCGCAAAACGATTAAAGAATGTGGTCC 1936
OY      430 GluValGlyLysLeuGlnAlaAlaAlaSpMetIleSerMetPheThrGlnPheTrpGln 449
       ||| ||| ||| ||| ||| ::| :|::| ||||| |||
Db      1937 AACGCGCTGAAAA--CAATGTGGCGCAAAACGCCAGCAAACTCTTTACCCAGGGGATGCAA 1993
OY      450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyLu 461
       ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db      1994 CGTATTACTAACGGCTGTGGTAATGTGGTAGCAAGATGGCGCTGCACAACGAATGCTT 2053
OY      462 SerAsnGluMet-----ThrGlnLysAlaThrLysLeuGlnAlaGlnIleLeuLys 478
       ||| ||||:::|||| ||| ::| ||||| ::|
Db      2054 AGTAAGAGAGCTGTGTAAGTAATACCTTAATTAAGTGGCGTTGGGCATGAAATCACGAAT 2113
OY      479 AlaTryAlaAlaIleSerGlyAlaIleAla 488
       ||| ::| ||| ::| ||| ::| |||
Db      2114 ACCGCAGCCGACGTCAGCGGTGTGTGGCC 2143

RESULT 11
US-08-961-527-71
; Sequence 71, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

Db	936	GCTCTGGGA-----GAGGCTCAGAGGCGAGCGATCTCTATGACGCCAGCATCAAA	986
OY	68	AspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaGly	87
Db	987	-----AGAGCGATACCGCCAGAGCTGTTATGACCGTCGGCC	1025
OY	88	LysGlnSerSerGlnSerGlnLysAla-----GlyAlaAspThrGlyVal	102
Db	1026	AAAAAAGTACGACGGCGGCAAAATAATGCAATGCTGAGACCAGCTGACCCGGGTAT	1085
OY	103	SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer	122
		:::	
		:::	
Db	1086	GCACAAAGCTGAAAGCC-----GCG	1103
OY	123	IleGlnGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla	142
Db	1104	GTAGAACAGCGCCGAAAAAGACGACAGCGCGAAGAGGCGCTTGATAGAGCCAGCAT	1163
OY	143	AlaGlnMetLys-----GluValGluAlaValAlaValAlaAlaLeuSer	157
Db	1164	GCACAGCGTTAAAGCAGCAGCAGACGCCAAGCCAAAGCCGAAAGCGATATCACTTGG	1223
OY	158	GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyAlaThr	177
Db	1224	ACCAAAATTCAGAGGAAGCGCTAATGCCCTCTCGAATCAGATGAGTTTCCAGGCTGACG	1283
OY	178	ProArg-----SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu	196
Db	1284	GATATCTGTCAAATATGTCGCCCGCTCATATGCCATGCGCATGTTATATGAGATTTGG	1343
OY	197	GlyValuAlaThrLysSerSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGln	216
Db	1344	GGCAAAATACGAGAAAGAACTGCAAAACGATCTTGGCTT-----	1385
OY	217	ThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGluThr	236
Db	1386	TTTCAAACGCTTGCAGGAGAGGCGCTCAGGG--GAGATGAAAGAAATACGCGTAATTC	1442
OY	237	GlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal	256
Db	1443	CAGGAA-----GAGACGCGGCAAGCCGAGAAACAGACACCGCATATGAGATGATTC	1493
OY	257	AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle	276
Db	1494	GGCAAAAGTCTCTCGCGCGCTG-----CTAACCTTGTCTGCGCTGTGCGCTGTT	1544
OY	277	PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaGlyAlaValGlyAlaAlaAla	296
Db	1545	TTTACCGGTGGCGCAGTGTGCGCTGCTGGTGGTGGACCTTGCGGTA-----	1592
OY	297	AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGln	316
Db	1593	-----ATGGGCGCCGATGAAAT-----	1610
OY	317	AlaValValGlnAlaValLysGlnAlaVal--IleThrAlaValAlaGlnAlaIleThr	335
Db	1611	-----GTGAAGCGCGGCGAGCGGGGGTGTGTTTATTCAGCGGCGCTAAAC	1655
OY	336	AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys	355
Db	1656	CCGATTATGGGCGCATGTG-----CTGAACCCGTTATATGAG	1691
OY	356	AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet	375
Db	1692	CTGATTTGGCAAGGCGATTTACCAAAAGCGCTG-----GAAGCATTTAGCGCTC	1736
OY	376	IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpVal	395
Db	1737	GATAGAAAAGCGCAGACATGGCAGCGAGCATTTGTGGTCCATCTGTCTCCCTATTGCC	1796
OY	396	ThrValGlyValGlyValValVal-----AlaAlaProAlaLeuGly	409
Db	1797	ATGCTATCGCGCTCATTTGTGTGTGTCGACAGTTCTGCGGAAGCGCGCGGCGCAAACTGGGT	1856

QY	410	lysG1Y11eMgc1nMeG1nleuSerc1umec1nG1nAsnY1AlaG1nPhc1nAlYs	429
Db	1857	AACGGCTGACGAAATG-----ATGGCGCAACGATTAGAGTGTGGCT	190
QY	430	GlUvAlc1Yls1eug1nAlAlAlAspMet1IeserwepheThrc1nPhetrc1n	449
Db	1905	AACGTGCTGAAA--CAGTTGCACAAACCGCAGCAAACTCTTACCCAGGGATGCA	196
QY	450	GlU-----AlAser1Ysl1eAlAser1Ysg1nThrc1nYglU	461
Db	1962	CcGATTACTAGCGCCCGGTATGTGCGTGCAGATGGCCCTGCAAACGAATGCCTTA	201
QY	462	SerAsn1umet-----Thrc1nYsAlArh1Yrs1eug1nYAlaG1n1leu1Ys	477
Db	2022	ACATAAGAGCTGTGATTAATACCTTAATATAAGGCGCTGGGCAAGACACGAAT	20
QY	479	AlATYrAlAlAlAlAserc1YAlAlAlAlAl	488
Db	2082	ACCGCAGCCAGTCAGCCGGGTGGCTTGC	211
RESULT 10			
US-08-591-079-7			
: Sequence 7, Application US/08591079			
: Patent No. 5972899			
: GENERAL INFORMATION:			
: APPLICANT: Zychlinsky, Arturo			
: APPLICANT: Chen, Yajing			
: TITLE OF INVENTION: Apoptosis Induced by Shigella Ipab			
: NUMBER OF SEQUENCES: 10			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: MORRISON & ROESTER			
: STREET: 2000 Pennsylvania Avenue, NW			
: CITY: Washington			
: STATE: DC			
: COUNTRY: USA			
: ZIP: 20006-1812			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.30			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/591,079			
: FILING DATE:			
: CLASSIFICATION: 435			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Iivnat, Shmuel			
: REGISTRATION NUMBER: 33,949			
: REFERENCE/DOCKET NUMBER: 15661-20017.00			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: (202) 887-1500			
: TELEFAX: (202) 887-0764			
: TELEX: 90-4030 MRSNFOERSMWH			
: INFORMATION FOR SEQ ID NO: 7:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 3622 base pairs			
: TYPE: nucleic acid			
: STRANDEDNESS: single			
: TOPOLOGY: linear			
: ORIGINAL SOURCE:			
: ORGANISM: Salmonella typhimurium			
: STRAIN: slbB			
: FEATURE:			
: NAME/KEY: CDS			
: LOCATION: 575..2356			
: OTHER INFORMATION: /product= "slbB"			
: OTHER INFORMATION: /gene= "slbB"			
US-08-591-079-7			
Alignment Scores:			
Pctd. No.: 5,26e-12 Length: 3622			
Score: 218.50 Matches: 116			

APPLICATION NUMBER: JP 106011/95  
 FILING DATE: 28-APR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Miller, Charles E.  
 REGISTRATION NUMBER: 24,576  
 REFERENCE/DOCKET NUMBER: 7426-043-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1296 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: Other nucleic acid; Synthetic DNA  
 US-08-809-326A-18

Alignment Scores:  
 Pred. No.: 1.44e-116 Length: 1296  
 Score: 1245.00 Matches: 259  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 52.55% Indels: 0  
 DB: 4 Gaps: 0

US-09-889-314-2 (1-496) x US-08-809-326A-18 (1-1296)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23  
 DB 484 ATGTCTATTTCATCTCTCTCAGAGACTGACAAATCAAAAAATATCATCTCATGCTCAGTTCG 543  
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
 DB 544 ACATGCACACCCAGGCGCTGCCCAACAAGATTAACCTGTGGCAACGAACGAAGCAA 603  
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
 DB 604 ATACAGCAACACGCTCAGGTAATAAACAACCTGACATGCAAGCAACGATTCCTGCTGGT 663  
 QY 64 AlaSerGlyLysAspLysThrSerSerThrLysThrLysGlnThrAlaProGlnGlnGly 83  
 DB 664 GCTTGGAAAGACAAACTCTCTGACTACAAAAACGAAGACAGCTCCACAAACAGGGA 723  
 QY 84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103  
 DB 724 GTTGCTGTGGGAAACAACTCTCAGAAAGTCAAAAGGAGGTGCTGATCTGAGATATCA 783  
 QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
 DB 784 GAGAGCGGCTGCTACTACACATCAAAATCTGCAACAAATCTGATGACAGACTCTATT 843  
 QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
 DB 844 GAAGAGCGAGCAAAAGTATGAGTCTACCTTAGTCACTTCAAAAGCCCTCAGTCCGGG 903  
 QY 144 GlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer 163  
 DB 904 CAATGTAAAGAGTCGACACGGTGTGTGCTGCCCTCAGAGGAAGATTGGGTTCC 963  
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183  
 DB 964 GCAAAATTTGGAACCTTAGCTCCCAAGCCCGGGGTGACCAAGATCAGAGGTATATC 1023  
 QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203  
 DB 1024 GAAATCGAGCTGCGCTCTTAAGCAATTCACACATTTGGAGAACCCCAAAATCTGCC 1083  
 QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln 223  
 DB 1084 TTATCTAATGATGACAGTACACAAAGCAAGACGACCAAAACAATTAACGTAGTCTAGAA 1143

QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlyTrpGlnGlnMetLysAlaAlaGln 243  
 DB 1144 AAGCAAGCATATAAATGATTAAGACGAGAGAAATACCAAGATGATGAAGCTCCGCA 1203  
 QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262  
 DB 1204 CAGAAGTCTAAAGATCTGAGAACAAATGATGATCTCAATACTGTGATGATCCGC 1260

# RESULT 9

US-08-591-079-9  
 Sequence 9, Application US/08591079  
 Patent No. 5972899  
 GENERAL INFORMATION:  
 APPLICANT: Zychlinsky, Arturo  
 TITLE OF INVENTION: Apoptosis Induced by Shigella IPAB  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 2000 Pennsylvania Avenue, NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/591,079  
 FILING DATE:

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Litvat, Shmuel  
 REGISTRATION NUMBER: 33,949  
 REFERENCE/DOCKET NUMBER: 15661-20017.00  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0764  
 TELEX: 90-4030 MRSNFOERSWSH  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5393 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: Salmonella typhi  
 STRAIN: Ty2  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 543..2324  
 OTHER INFORMATION: /gene="sfpB"  
 US-08-591-079-9

Alignment Scores:  
 Pred. No.: 6.3e-12 Length: 5393  
 Score: 220.50 Matches: 116  
 Percent Similarity: 41.02% Conservative: 85  
 Best Local Similarity: 23.67% Mismatches: 182  
 Query Match: 9.31% Indels: 107  
 DB: 2 Gaps: 20

US-09-889-314-2 (1-496) x US-08-591-079-9 (1-5393)

QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIleGlnGlnThr 47  
 DB 876 CAGCGCATGTTGATGACAAAGAGATGGGATTCAGATCCGAAGCAATTCACAGAG 935  
 QY 48 ArgGlnLysAsnThrGlnMetGlnSerAspAlaThrIleLeuAlaGlyLysSerGlyLys 67

QY	4	MetSerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGluValLeu	23
Db	1	ATGTCATTATTCATCTTCTTCAGACCTGCACAAACCAAAAAAATATCATGTCTCCAGTTCTG	60
QY	24	ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGln	43
Db	61	ACATGACACCCCGAGGGGGTCCCAACAAAGATATACGTCTCTGCACCAACCAAGGAA	120
QY	44	IleGlnGlnThrArgGlnGlnIlyAsnThrGluMetGluSerSerSplaThrIleLeuIleagly	63
Db	121	ATACACCAACACGTCAGGGTAAAAACACTGAGTGGAAAGGCATGTCCACTTTGCTGGT	180
QY	64	AlaSerGlyLysAspLysThrSerSerThrThrIlyThrGluThrAlaProGlnGlnIly	83
Db	181	GCCTTCGGAAAAACAAACTTCTCCGACGTCACAAAACAGAACAGCTCCACAAACGGGA	240
QY	84	ValAlaAlaGlyLysGluSerSerGluSerGlnLysIleaglyAlaAspThrGlyValSer	103
Db	241	GTTGCTGCTGGGAAGAATCTTCACAAAGTCAAAAGCAGGTGCTGTACTGCAGATATCA	300
QY	104	GlyValAlaAlaThrThrThrLaserAsnThrAlaThrLysIleAlaMetGlnThrSerIle	123
Db	301	GGACCGGCTCTACTACAGCATCAATATCTGCACAAAAAATGGTATATGCACACCTCTATT	360
QY	124	GluGluAlaSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla	143
Db	361	GAAAGCGGACCAAAAGTAGAGACTCTACCTTAAGATCACTTCAAAAGCCTCAGTGGCGG	420
QY	144	GlnMetLysGluValGluAlaValAlaValAlaLeuSerGlyLysSerSerGlySer	160

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fennie E. Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:



```

STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995

ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA

US-08-809-326A-7

Alignment Scores:
Pred. No.: 6,3e-117 length: 777
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.55% Indels: 0
DB: 4 Gaps: 0

US-09-889-314-2 (1-496) x US-08-809-326A-7 (1-777)

QY 4 MetSerIleSerSerSerSerglyProAspAsnGlnIlysaSnIleMetSergInValIeu 23
Db 1 ATGCTATTTCATCTCTCTTCAGACCTGCAGCAATCAATCAAAAATAATATCATGTCTCAAGTTCTG 60

QY 24 ThrSerThrProGlnGlyValProGlnGlnAspIlyseuSergIlysaGngIuThrIlysgln 43
Db 61 ACATCGACACCCCGAGGCGTCCGCCCAACAAGATTAAGCTGTGGCAACGAAACGAACAA 120

QY 44 IleGlnGlnThrArgGlnGlyIlysaSnThrGluMetGluSerAspAlaThrIleAlagly 63
Db 121 ATACAGCAACACCTCAGGCGTAATAAACACATGAGATGGAAGACCATGCTATTTCGTGCT 180

QY 64 AlaSerGlyIlysaSplysThrSerSerThrThrIlysthrGluThrAlaProGlnGlnGly 83
Db 181 GCTTCTGGAAAAAGACAAACCTTCTCGACTTACAAAACAGAAACAGCTCCACAAACAGGA 240

```

```

101  Oy      84      ValAlaAlaGlyLysGluSerSerGluSerLysLysAlaGlyAlaAspPheGlyValSer 103
102  Db      241      GTTGCTGCTGGGAAAGAAATCCAGAAAGTCAAAAAGCGAGGTGCTGATACGAGATCA 300
103  Oy      104      GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
104  Db      301      GGAGCGGCTGCTACTACGACATCAAAATCTGCACAAAATAATTGCTATGACAGCTTATT 360
105  Oy      124      GluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAla 143
106  Db      361      GAAGAGCGGACCAAAAGATGAGCTACTAGAGTCACTTCAAGGCTCAGTGGCCGC 420
107  Oy      144      GlnMetLysGluValGluAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
108  Db      421      CAAATGAAAGAAAGTCGAAGCGGTGTGTTGCTCTGCCCTCTCAGGGAAAAAGTTCCGGATTCC 480
109  Oy      164      AlAlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
110  Db      481      GCMAAATTGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATTCAGAGTTATC 540
111  Oy      184      GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203
112  Db      541      GAAATCGGACTCGCGCTGTGTAAGCAATTCAGACATTCGGGAGAAGCCACAATAATCTGCC 600
113  Oy      204      LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
114  Db      601      TTATCTTAATCTATGCAAGTACACAAAGCACACAGCCAAACAAAATAAATCTGAGCTTAGAA 660
115  Oy      224      LysGlnAlaIleLysIleAspLysGluArgGluGluGlyTrpGlnGluMetLysAlaGlu 243
116  Db      661      AAGCAAGCGATAAAAATGATGATAAAGACAGACAGATATCCAAAGATGAAGGCTGCCAA 720
117  Oy      244      GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAla 262
118  Db      721      CAGAAGCTAAAGATCTCGAAGAAACAAATGATGATCTGCACTACTGATGATGATCGCG 777
119  RESULT 7
120  US-08-809-326A-4
121  ; Sequence 4, Application US/08809326A
122  ; Patent No. 6165478
123  ; GENERAL INFORMATION:
124  ; APPLICANT: Izutsu, Hiroshi
125  ; APPLICANT: Obata, Kazuhiko
126  ; APPLICANT: Matsumoto, Akira
127  ; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
128  ; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, AND TRANSFORMA
129  ; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
130  ; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
131  ; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUR
132  ; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DIAGNOSIS OF CHLAMYDIA
133  ; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
134  ; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
135  ; TITLE OF INVENTION: PNEUMONIAE GENE
136  ; NUMBER OF SEQUENCES: 31
137  ; CORRESPONDENCE ADDRESS:
138  ; ADDRESSEE: Pennie & Edmonds LLP
139  ; STREET: 1155 Avenue of the Americas
140  ; CITY: New York
141  ; STATE: New York
142  ; COUNTRY: USA
143  ; ZIP: 10036-2711
144  ; COMPUTER READABLE FORM:
145  ; MEDIUM TYPE: Floppy disk
146  ; COMPUTER: IBM PC compatible
147  ; OPERATING SYSTEM: PC-DOS/MS-DOS
148  ; SOFTWARE: PatentIn Release #1.0, Version #1.30
149  ; CURRENT APPLICATION DATA:
150  ; APPLICATION NUMBER: US/08/809,326A
151  ; FILING DATE: 19-MAR-1997
152  ; CLASSIFICATION: 435
153  ; PRIOR APPLICATION DATA:
154  ; APPLICATION NUMBER: JP 224711/94

```

CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/809,326A  
 FILING DATE: 19-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 224711/94  
 FILING DATE: 20-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106006/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106008/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106009/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106010/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106011/95  
 FILING DATE: 28-APR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Miller, Charles E.  
 REGISTRATION NUMBER: 24,576  
 REFERENCE/DOCKET NUMBER: 7426-043-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1048 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: Genomic DNA  
 ORIGINAL SOURCE:  
 ORGANISM: Chlamydia pneumoniae  
 STRAIN: YK-41  
 IMMEDIATE SOURCE:  
 CLONE: 53-35  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 236 to 1012  
 IDENTIFICATION METHOD: P  
 US-08-809-326A-9  
 Alignment Scores:  
 Pred. No.: 1,95e-118 Length: 1048  
 Score: 1262.00 Matches: 262  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 53,278 Indels: 0  
 Gaps: 0  
 DB:  
 US-09-889-314-2 (1-496) x US-08-809-326A-9 (1-1048)  
 QY 1 AsphrsmksetserleserSerserGlyProaspasnGlnlysasnilekmetSer 20  
 Db 227 GATCAACAACATGCTATTCTCTTTCAGACGCTGACAAATCAAAATATCATGCTC 286  
 QY 21 GlnValleuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40

Db 287 CAAGTTCGACATCGACACCCAGGCGCTGCCCAACAAGATAGTGTCTGCAACGAA 346  
 QY 41 ThrlysglnileglnGlnThrArglnGlyLysasnThrGluMetGluSerAspAlaThr 60  
 Db 347 ACGAAGCAAAATACAGCAAAACACGTCAGGCTAAACACATGAGATGGAAGGATGCCACT 406  
 QY 61 lleAlaGlyAlaSerGlyLysAspLysThrSerThrThrLysThrGluThrAlaPro 80  
 Db 407 ATTCCTGGTGTCTTCGAAAAGCAAACTTCCTGCATACAAAAACGAAACAGCTTCCA 466  
 QY 81 GlnGlnGlyValAlaAlaGlnLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100  
 Db 467 CACACAGGAGTTCTCTGCGAAGATCTCGAAGAGTCAAAAGCGGTGCTGATACT 526  
 QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120  
 Db 527 GGAGTATCAGACAGCGGCTGCTACTACAGATCAAAATACGCAACAAAATTGCTATGCG 586  
 QY 121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu 140  
 Db 587 ACCTTAATGAAGGCGGACCAAAAGTATCGAGTCTTACCTTAGACATTCAAAAGCCTC 646  
 QY 141 SerAlaAlaGlnMetLysGluValGluAlaValAlaAlaLeuSerGlyLysSer 160  
 Db 647 AGTCCCGCCCAAAATGAAGAAGTCGAAGCGGTGTGTGCTGCTCCACAGGAAAGT 706  
 QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180  
 Db 707 TCGGATCCGCAAAATTGGAAACCTGACCTCCCAAGCCGGGTGACACCAAGATCA 766  
 QY 181 GlnValIleGlnIleGlyLeuAlaAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThr 200  
 Db 767 GAGGTATCGAAATCGGACTCGCGCTGTCAAAACATTGACATTTGGAGAGGCCACA 826  
 QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220  
 Db 827 AAATCTGCTTATCTTAATCTATGCAAGTACACACACAGACAGACCAAAATAAATACTA 886  
 QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlyTrpGlnGlnMetLys 240  
 Db 887 GGCTTAGAAAACCAAGCCATATAAATCGATTAACAGCAAGAAATGCCAAGAGATGAAG 946  
 QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260  
 Db 947 GCTGCCGACAGAGCTTAAGATCTCGAAGGAACAATGATCTGCATTAATCTGATG 1006  
 QY 261 lleAla 262  
 Db 1007 ATCGCG 1012  
 RESULT 6  
 US-08-809-326A-7  
 ; Sequence 7, Application US/08809326A  
 ; Patent No. 6165478  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Izutsu, Hiroshi  
 ; APPLICANT: Obata, Kazuhiko  
 ; APPLICANT: Matsunoto, Akira  
 ; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
 ; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 ; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
 ; TITLE OF INVENTION: CONTRAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
 ; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM  
 ; TITLE OF INVENTION: OF ANTIBODY, METHOD AND REAGENTS FOR DIAGNOSIS OF CHLAMYDIA  
 ; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
 ; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
 ; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York

US-08-809-326A-25

## Alignment Scores:

Pred. No.:	1-4e-117	Length:	5438
Score:	1265.00	Matches:	312
Percent Similarity:	62.39%	Conservative:	33
Best Local Similarity:	56.42%	Mismatches:	75
Query Match:	53.40%	Indels:	135
DB:	4	Gaps:	10

US-09-889-314-2 (1-496) x US-08-809-326A-25 (1-5438)

```

OY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
DB 540 ATGTCTATTTCATCTTCTTCAGAGACCTGCATCAAAAATAATCATGTCTCAAGTTCCG 599
OY 24 ThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
DB 600 ACATGACACCCCGAGGCGCTGCCCAACAAGATTAAGCTCTGGCCAAACGAAGCAA 659
OY 44 IleGlnIleThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
DB 660 ATACAGCAAAACGTCAGGCTAATAAACACTGAGATGGAAGCGATGCCACTATGTGCTGT 719
OY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrLysThrLysThrLysThrLys 83
DB 720 GCTTCGGAAAGACAAACCTTCCTCGACTACAAAAACGAAGACGCTCCACAGCGGA 779
OY 84 ValAlaIleAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 780 GTTGTGCTGGGAAAGAACTTCAGAAAGTCAAAAGGCGTGTGATCTGAGATCA 839
OY 104 GlyAlaAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetIleThrSerIle 123
DB 840 GGAGCGGCTGCTACTACACCACTCAAAATCTGCTATGAGACCTTAT 899
OY 124 GlnGlnAlaSerLysSerMetGlnSerThrLysGlnSerLysGlnSerLysSerAlaAla 143
DB 900 GAAGAGGCGAGCAAAAGTATGAGTCTACCTTATGAGTCACTTCAAAAGCCTTCAGTCCGCG 959
OY 144 GlnMetLysGlnValGlnAlaValValAlaAlaIleLeuSerGlyLysSerSerGlySer 163
DB 960 CAAATGAAGAAAGTCGAACGGTGTGTGTGCTGCGCTTCAGGAAAGTTCGGGTTCC 1019
OY 164 AlaLysLeuGlnIleThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
DB 1020 GCAAAATTTGAAACACCTAGCTCCCAAGCCCGGGTACACCAAGATCAGAGTTATC 1079
OY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLysGlnGlyAlaIleThrLysSerAla 203
DB 1080 GAATCGGACCTGGCGCTCTTAAGCAATTCAGACATTTGGAGAAAGCCCAAAATCTGCC 1139
OY 204 LeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnIleGlu 223
DB 1140 TTATCTAATCTACCAAGTCAACAGCAAGACAGCAACCAAACTAAGTCTAGGA 1199
OY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleGlnThrLysGlnMetLysAlaIleGlu 243
DB 1200 AAGCAAGCATTAATAATTCATTAAGACAGAGATACCAAGATGAGAGGCTGCCGA 1259
OY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 1260 CAGAAGTCTTAAGATCTCAAGAGCAATGATGATCTCAATAGTGTGATGATCGGAA- 1318
OY 264 SerValAlaIleThrVal-----IleSerIleValAla 274
DB 1319 GGGGTTGCAATTGGCATGGGGCCCTTAATTAATTAATCGAGAGATCCAGATCTATG 1378
OY 275 AlaIlePheThrCysGlyLysLeuAlaGlyLeu-----AlaAlaGlyAlaVal 292
DB 1379 ATGATCTCTAGCGCGAGCATCGTGGCGGATCACCAGCGGCCACAGTCCGTTGCT 1438
OY 293 GlyAlaAla----- 295

```

```

DB 1439 GCGCGCTTA-TATCGCGCATCATCCGATGGGGAAGATCGGGCTGCCACTTCGGGCTCAT 1497
OY 296 -----AlaIleGlyAlaAlaGlyAla 303
DB 1498 GAGCGCTGTGTTGCGCGGTGATGTGGCAGCGCCGTGGCGGGGACTGTGGGCGCC 1557
OY 304 -----AlaAlaAla----- 306
DB 1558 ATCTCTTGACATGACACCATCTCTTGGCGGGGTGCTCAACGGCTCACTACTACTG 1617
OY 306 ----- 306
DB 1618 GGCTGCTTCTTAATGACAGAGTCGATTAAGGAGAGCGTCGACGATGCCCTTGAGAGCC 1677
OY 307 -----ThrValAlaThrGlnIle 313
DB 1678 TTCAACCCAGTCAGCTCTTCGCGTGGCGCGGACATGATGCTGCGCCACTTATG 1737
OY 314 ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAla 333
DB 1738 ACTGCTTCTTATCATGCACTGTAAGACAG----- 1770
OY 334 IleThrAlaAlaIle-----LysAlaAlaValLys 343
DB 1771 GTCGCGGACGCGCTGCGGTATTTTCGGCAGAGACCGCTTCGCTGGAGCGCGCATG 1830
OY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
DB 1831 ATCGGCTGCTGCTTCGCGTATTCGGAATCTTCACGCGCTTCGCAAGCCTTCCTCACT 1890
OY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspPheProLysLeu 383
DB 1891 GGTCCCGCC-----ACCAACGTTTCGCCGAGAA 1920
OY 384 SerLysValIleSerSerLeu----- 390
DB 1921 CAGCGCATATTCGCGGATGCGCGCGCGCGCTGAGCTGCTGCTGCGCTTCGCG 1980
OY 391 ThrSerLysThrValThrValGlyValAlaValAlaAlaProAlaLeuGlnLys 410
DB 1981 ACGGAGGCTGGATGCTTCCTCCCATTTATGATTTCTTCGCTTCGCGCGCGCGCATG 2040
OY 411 GlyIleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPheGlnLysGlu 430
DB 2041 CCGCGGTTCCAGGCCATGCTGTCCAGCGCAGGTAGATGACAGCATCAGGAGACCTTCAA 2100
OY 431 ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
DB 2101 ---GGATCGCTCGCGGCTCTTACCAAGCTTAACCTGATC 2136

RESULT 5
US-08-809-326A-9
; Sequence 9, Application US/08809326A
; Patent No. 6165478
; GENERAL INFORMATION:
; APPLICANT: Iizutsu, Hiroshi
; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds LLP
; STREET: 1155 Avenue of the Americas

```

Db 1231 TCGGTTCGGCAAAATTTGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290  
 Oy 181 GlnValIleGluIleGlyLeuAlaLeuAlaValAlaIleGlnThrLeuGlyValaThr 200  
 Db 1291 GAGGTTATGCAATTCGACTCGCGCTTGCTAAAGCAATTCAGACATTTGGAGAACCCACA 1350  
 Oy 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrSerLysLeu 220  
 Db 1351 AAATCTCTTATCTAATCACTATGCAAGTACCAAGACCAAGACCAAGCAAAATTAACCTA 1410  
 Oy 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysLeu 240  
 Db 1411 GGCTTACAAAAGCAAGCGATTAATAATGATAAAGCAAGAAAGATACCAAGATGAAG 1470  
 Oy 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetLysPheValAsnThrValMet 260  
 Db 1471 GCTCCCGACAGAGAGTCTAAAGATCTGAAGCAATGATCTGTCATCTACTGTGATG 1530  
 Oy 261 IleAlaValSerValAlaIleThrVal-----LysSer 271  
 Db 1531 ATCGGAA-GGGGTTCGATTCGATGGGGGCGCTTAATTAACTCGAGAGATCCAG 1589  
 Oy 272 IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly 289  
 Db 1590 ATCTAATGATGATCTCTACGCGGACGATCGTGGCGGCGCATCCGCGCCACAGGT 1649  
 Oy 290 AlaAlaValGlyAlaAla----- 295  
 Db 1650 GCGGTTCGGCGCCTA-TATCGCCGACATACCGATGGGAGAAATCGGGCTGCCACTT 1708  
 Oy 296 -----AlaAlaGlyVala 300  
 Db 1709 CCGGCTCATGACGCTTGTTCGGGTGATGGTGGCAGGCGCGGCGGAGCTG 1768  
 Oy 301 AlaGlyAla-----AlaAlaAla----- 306  
 Db 1769 TTGGGCGCATCTCTTGATGATGACCATCTCTTGCGGCGGCGGTCTCAAGGCTCAAC 1828  
 Oy 306 ----- 306  
 Db 1829 CTACTAGTGGGCTCTCTTAATGACGAGATGCGATTAAGGAGAGCGTCGACCGATGCC 1888  
 Oy 307 -----ThrValAla 310  
 Db 1889 TTGAGAGCTTCAACCCAGTACGCTCTTCGGTGGCGCGGCGGCGATGACTCTGTCGCG 1948  
 Oy 311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330  
 Db 1949 GCACCTATGACTGTCTTCTTATCATGCAACTCTGAGACAG----- 1990  
 Oy 331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340  
 Db 1991 -----TGCCCGCAGCGCTCTGGTCAATTTGGCGGAGACAGCCCTTGCTGGAGC 2041  
 Oy 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360  
 Db 2042 GCGACGATGATGCGCTGTGCTGCGGTATGCGATCTTGACGCGCTCGCTCAAGCC 2101  
 Oy 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380  
 Db 2102 TTTCGCACTGCTCCGCC-----ACCAACGTTTC 2131  
 Oy 381 ProLysLeuSerLysValIleSerSerLeu----- 390  
 Db 2132 GCGGAGAACGAGCGCATTAATGCGCGCATGGCGGCGCGGTGGCTAGCTTCTGCTG 2191  
 Oy 391 -----ThrSerLysTrpValThrValGlyValGlyValValAlaAlaAlaProAla 407  
 Db 2192 GCGTTCGAGACGAGCGGTGATGCTTCCCATTAAGATCTTCGCTCGGCGCGC 2251  
 Oy 408 LeuGlyLysGlyIleMetGlnMetLysSerGlnMetGlnGlnAsnValAlaGlnPhe 427  
 Db 2252 ATCGGATGCCCCGCTTGACGCAATGCTGTCCAGGCGATGATGACGACACATCAGGGA 2311

Oy 428 GlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443  
 Db 2312 CAGCTTCAA---GGATCGCTCGCGCTCTTACCAAGCTTAATTCGATC 2356  
 RESULT 4  
 US-08-809-326A-25  
 Sequence 25, Application US/08809326A  
 Patent No. 6165478  
 GENERAL INFORMATION:  
 APPLICANT: Izutsu, Hiroshi  
 APPLICANT: Matsumoto, Akira  
 APPLICANT: Obara, Kazuhiko  
 TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING  
 THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
 TION OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
 TION OF INVENTION: CONTRAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
 F ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
 MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
 REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 PNEUMONIAE GENE  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSER: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/809,326A  
 FILING DATE: 19-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 224711/94  
 FILING DATE: 20-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106006/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106008/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106009/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106010/95  
 FILING DATE: 28-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 106011/95  
 FILING DATE: 28-APR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Miller, Charles E.  
 REGISTRATION NUMBER: 24,576  
 REFERENCE/DOCKET NUMBER: 7426-043-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 INFORMATION FOR SEO ID NO. 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5438 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: Other nucleic acid; Plasmid

Oy	344	SergilyIeIySAlaPheIleIySthrIeUValIySAlaIleAlaIleSerIyS	363
Db	1504	TCTGCAATAAAAAGCATTTATCAAAACTTTAGTCAAAGCGATTGCCAAAGCCATTTCTAAA	1563
Oy	364	GlyIleSerIySValPheAlaIySgLythrGlnMetIleAlaIySAsnPheProIySLeu	383
Db	1564	GGAATCCTTAAGGTTTTCGCTAAGGAGCACTAAATGATTGGGAAGAACTTCCCAAGCTC	1623
Oy	384	SerIySValIleSerSerIeUthrSerIyStrpValIthrValGlyValGlyValVal	403
Db	1624	TGGAAGATCATCTGCTCTTTCACAGTAATGGGACAGGTGGGGTGGGTGTAATT	1683
Oy	404	AlaAlaProAlaIeUgLyIySgLyIleMetGlnMetGlnLeuSerGlyMetGlnGlnAsn	423
Db	1684	GCGGGCCCTGCTCTCGGTAAAGGATTATGCAAATGCAAGCTCTCGCAATGCACAAATAC	1743
Oy	424	ValAlaGlnPheGlnIySgLyValGlyIySLeuGlnAlaIleAlaIleAspMetIleSerMet	443
Db	1744	GTCGCTCAATTTCAAAAAGAGCTCGAAAACTGCGAGCGCGCTGATATATTTCTATG	1803
Oy	444	PheThrGlnPheTrpGlnAlaIleSerIySAlaSerIySAlaSerIySglnThrGlnIySerAsn	463
Db	1804	TTCACCTCAATTTTGGCAACAGCAAGTAATATGCTCTAAACAACACAGGCGAGCTATAT	1863
Oy	464	GlnMetThrGlnIySAlaThrIySLeuGlyAlaGlnIleLeuIySAlaTyAlaAlaIle	483
Db	1864	GAAATGACCTCAAAAAGCTACCAAGCTGGCGGCTCAAAATCTTTAAAGCGTATGCGCCATTC	1923
Oy	484	SerGlyAlaIleAlaGlyAla 490	
Db	1924	AGCGAGGCATCGCTGGCGCA 1944	

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  JP 106008/95
3      FILING DATE: 28-APR-1995
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  JP 106009/95
6      FILING DATE: 28-APR-1995
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER:  JP 106010/95
9      FILING DATE: 28-APR-1995
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER:  JP 106011/95
12     FILING DATE: 28-APR-1995
13     ATTORNEY/AGENT INFORMATION:
14     NAME:  Miller, Charles E.
15     REGISTRATION NUMBER:  24,576
16     REFERENCE/DOCKET NUMBER:  7426-043-999
17     TELECOMMUNICATION INFORMATION:
18     TELEPHONE:  (212) 869-9090
19     TELEFAX:  (212) 869-8864/9741
20     TELEX:  66141 PENNIE
21     INFORMATION FOR SEQ ID NO:  10:
22     SEQUENCE CHARACTERISTICS:
23     LENGTH:  5658 base pairs
24     TYPE:  nucleic acid
25     STRANDEDNESS:  double
26     TOPOLOGY:  linear
27     MOLBUCE TYPE:  Other nucleic acid; Plasmid
28     US-08-809-326A-10

```

Alignment Scores:	
Pred. No.:	2,85e-119
Score:	1282.00
Percent Similarity:	62.59%
Best Local Similarity:	56.65%
Query Match:	54.12%
DB:	4
	Gaps:
	10
	135
	Indels:
	73
	Mismatches:
	315
	Conservative:
	315
	Matches:
	5655
	Length:

US-09-889-314-2 (1-496) x US-08-809-326A-10 (1-5658)

1 APPLICANT: Matsumoto, Akira  
 2 TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING  
 3 TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 4 TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS  
 5 TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
 6 TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
 7 TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
 8 TITLE OF INVENTION: PNEUMONIAE INFECTIONS. PROBES AND PRIMERS FOR DETECTION AND/OR  
 9 TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
 10 TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 11 TITLE OF INVENTION: PNEUMONIAE GENE  
 12 NUMBER OF SEQUENCES: 31  
 13 CORRESPONDENCE ADDRESS:  
 14 ADDRESSEE: Pennie & Edmonds LLP  
 15 STREET: 1155 Avenue of the Americas  
 16 CITY: New York  
 17 STATE: New York  
 18 COUNTRY: USA  
 19 ZIP: 10036-2711  
 20 COMPUTER READABLE FORM:  
 21 MEDIUM TYPE: Floppy disk  
 22 COMPUTER: IBM PC compatible  
 23 OPERATING SYSTEM: PC-DOS/MS-DOS  
 24 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 25 CURRENT APPLICATION DATA:  
 26 APPLICATION NUMBER: US/08/809,326A  
 27 FILING DATE: 19-MAR-1997  
 28 CLASSIFICATION: 435  
 29 PRIOR APPLICATION DATA:  
 30 APPLICATION NUMBER: JP 224711/94  
 31 FILING DATE: 20-SEP-1994  
 32 PRIOR APPLICATION DATA:  
 33 APPLICATION NUMBER: JP 106006/95  
 34 FILING DATE: 28-APR-1995

Oy	1	AsHrshnmetSerIleSerSerSerGlyProAspGlnIlyAsnIleMetSer	20
Db	751	GATCAACACATGCTATTATCTCTTCTTCAGCACTGCAATCAAAAAAATATCATGCT	810
Oy	21	GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspIlySerGlyAsnGln	40
Db	811	CAAGTTCTGCATCGACACCCCAAGGGGCTGCCCAACAAGATAGCTGTCTGGCAACGA	870
Oy	41	ThrIysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr	60
Db	871	ACGAGACAAATACACGCAACACGCTCAGGGTAAAAAACACTGAGTGGAAAGCATCATCCACT	930
Oy	61	IleIleAlaGlyAlaSerGlyLysAspIlySerThrSerSerThrThrIysThrGlnIleAlaPro	80
Db	931	ATTCTGGTGGCTTGTGGAAAGACAAAAACTCTCTGCAGCTACAAAACAGAAACACTCTCA	990
Oy	81	GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnIlyAspAlaGlyAlaAspThr	100
Db	991	CAACAGGAGGTGCTGCTGGGAAAGAAATCCTCGAAAGTCAAAAGCAGCTGTATACT	1050
Oy	101	GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln	120
Db	1051	GGAGATCATCAGGAGCGCTGCTACTACAGCATCAAAATACGCAACAAAATTTGCTATGCGAG	1110
Oy	121	ThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeu	140
Db	1111	ACCTCATATGAGAGGGGACCAAAAGATGGAATCTTACTTATGAGTCACTTCAAAAGCCTC	1170
Oy	141	SerAlaAlaGlnMetCysGluValGluAlaAlaValAlaValAlaAlaLeuSerGlyLysSer	160
Db	1171	AGTCCGGCGCAATGAAAGAAGTCGAAGCGGTGTGTGTGCTGCTCTCAGGGAAGT	1230
Oy	161	SerGlySerAlaLysLeuGlnIleThrProGlnLeuProLysProGlyValThrProArgSer	180

```

? Sequence 17 Application US/08809326A
? Patent No. 6165478
? GENERAL INFORMATION:
? APPLICANT: Izutsu, Hiroshi
? APPLICANT: Obata, Kazuhiko
? APPLICANT: Matsumoto, Akira
? TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
? TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS TRANSFORMANTS
? TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, METHOD FOR PRODUCTION OF
? TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
? TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
? TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
? TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
? TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
? TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
? NUMBER OF SEQUENCES: 31
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pennile & Edmonds LLP
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/809,326A
? FILING DATE: 19-MAR-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA: JP 224711/94
? APPLICATION NUMBER: JP 224711/94
? FILING DATE: 20-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106006/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106008/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106009/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106010/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106011/95
? FILING DATE: 28-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Miller, Charles E.
? REGISTRATION NUMBER: 24,576
? REFERENCE/DOCKET NUMBER: 7426-043-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-8864/9741
? TELEX: 66141 PENNILE
? INFORMATION FOR SEQ. ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1947 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: other nucleic acid; Synthetic DNA
? US-08-809-326A-17
?
? Alignment scores:
? Pred. No.: 1.28e-224 Length: 1947
? Score: 2316.00 Matches: 487
? Percent Similarity: 100.00% Conservative: 0
? Best Local Similarity: 100.00% Mismatches: 0

```

Query Match:	97.76%	Indels:	0
DB:	4	Gaps:	0
US-09-889-314-2 (1-496) x US-08-809-3264-17 (1-1947)			
4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu	23		
484 ATGTATTATTCATCTTCTTCAGGACCTCGAATAATCAAAAATATCATGTCTCAAGTTCG	543		
24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln	43		
544 ACATCGACACCCCGAGGCGCTCCCAACAAGATAGCTGTCTGGCAACGAAGCA	603		
44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetLeuSerAspIleThrIleArgIle	63		
604 ATACAGCAACACGCTCAGGAGAAAACCTGAGATGAGAAAGCCATGTCATTTGCTGCT	663		
64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly	83		
664 GCTTCGGAAAAAGACAAAACCTCCCTCGACTCAAAAACAGAACAGCTCCACACAGGA	723		
84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyAlaSer	103		
724 GTTGCTGCTGGGAAACAAATCTCAAAAAGTCAAAAGGAGGTGCTGATCTAGAGATCA	783		
104 GlyAlaAlaAlaThrThrAlaSerSerThrAlaThrLysIleAlaMetGlnThrSerIle	123		
784 GGACCGCTCTCTACTACAGCTCAATACTGTCACAAAATTTGCTATGACAGCTCTATT	843		
124 GlnGlnLysLeuSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla	143		
844 GAAAGCGCGAACAAGATATGAGTGTACTCTTAGAGTCACTTCACAAAGCTCATAGTCCGCG	903		
144 GlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer	163		
904 CAAATGGAAGCAGTCGAAAGCGGTGTGTTGCTGGCCCTCTCAGGAAAAAGTTCGGGTTC	963		
164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle	183		
964 GCMAAATTTGGAACACTGAGCTCCCAACCCGGGGTGAACCAAGATCAGAGTTATC	1023		
184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSerAla	203		
1024 GAAATCGGACCTCGCGCTTGCTAAAGCAATTCACAGCATTTGGAGAACCCCAAAAATCTGCC	1083		
204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGlyLeu	223		
1084 TTATCTAATCTATGACAGTACCAAGACCAACCAACCAAAATTAATAGTCTACGA	1143		
224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnMetLysAlaAlaGln	243		
1144 AACCAAGCATTAATAATCGATTAAGACGAAACGAATAACCAAGATGAAGGCTGCCGA	1203		
244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal	263		
1204 CAAAGCTTAAAGATCTCGAAGGAACAATGATCTGTCAATACTGTGATGATCGCGGT	1263		
264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu	283		
1264 TCTGTTCCTTACAGTATTTCTATTTGTTGCTGATTTTACATCGCGAGCTGCACCTC	1323		
284 AlaGlyLeuAlaIleGlyAlaAlaValGlyAlaAlaAlaIleGlyAlaIleGlyAla	303		
1324 GCTGAGATGCTCGGGAGCTGCTGTAGGTGACAGGGAGCTGGAGGGTGCACGGAGCT	1383		
304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaAlaValAlaGlnAlaValLys	323		
1384 GCTCCGCCAACACAGGTAGGACAAACAATTCAGTTCAAGCTGTGTCCAGCGGTGAA	1443		
324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys	343		
1444 CAAGCTGTATACAGCTGTACAGCAAGCATCCGCGGCTATATAAGCGGCTCTCAAA	1503		

APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106011/95  
FILING DATE: 28-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Charles E.  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 7426-043-999  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ. ID NO.: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Other nucleic acid; Synthetic DNA  
US-08-809-326A-3

Alignment Scores:  
Pred. No.: 8,066-225 Length: 1464  
Score: 2316.00 Matches: 487  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 97.76% Indels: 0  
DB: 4 Gaps: 0

US-09-889-314-2 (1-496) x US-08-809-326A-3 (1-1464)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23  
DB 1 ATGTCTATTTCATCTCTTCAGAGCCTGACCAATCAAAAAATATCATGTCTCAAGTTCTG 60  
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43  
DB 61 ACATCGACACCCGAGGCGTCCCAACAAGATTAAGCTGTCTGGCAACGAAAGCAAA 120  
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
DB 121 ATTCACACAAACAGTCAGGTTAAACACACTGAGATGGAAGCGATCCACTATTGCTGG 180  
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83  
DB 181 GCTTTCGGAAAAACAAACTTCTCGACTACAAAAACAAACAAACACTCCACAAACAGGGA 240  
QY 84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrThrGlyValSer 103  
DB 241 GTTGCTGCTGGGAAATCTCTAGAAAGTCAAAGGACGAGTCTGTACTGAGATATCA 300  
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123  
DB 301 GGAGCGGCTGCTACTACAGATCAAAATFACTGCAACAAAATTTGCTATGACAGCTCTATT 360  
QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143  
DB 361 GAAGAGGCGAGCAAAAGTATGAGTCTACCTTACCTTCAAGCCCTCACTGCGCG 420  
QY 144 GlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer 163

DB 421 CAAATGAAGAGCTGAACGGTGTGTGCTGCCCTCTCAGGAAAAAGTTCCGGGTTC 480  
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183  
DB 481 GCAAAATTGGAAACACCTGAGCTCCCAAGCCGGGGTTCACCAAGATCAAGAGTTATTC 540  
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnValAlaThrLysSerAla 203  
DB 541 GAATTCGAGCTGCGCTTCTTAAGCAATTCAGACTTGGAGAAAGCCAAATATCTGCC 600  
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGln 223  
DB 601 TTATCTAATTAAGCAAGTACACAGACAGACAGACCAACCAATTAATTAAGTCTAGAA 660  
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlyThrGlnGlnMetLysAlaAlaGlu 243  
DB 661 AAGCAAGCATTAATAATCATTAAGAACAGAACAGAACATCCAGATGAAAGCTCCGCA 720  
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263  
DB 721 CAGAGCTTAAGATCTGAGAGAACAAATGATAGTCAATACTGTGATGATCGCGTT 780  
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283  
DB 781 TCTGTTGCCATTACAGATTATTTCTATTGTTGCTGCTATTTTACATCGGAGCTGAGTTC 840  
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyLysAlaAlaGlyAla 303  
DB 841 CGTGACCTGCTGCGGAGAGCTCTGTAGGTGCAGCGGACGTGAGGTGCACAGAGCT 900  
QY 304 AlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323  
DB 901 GCTGCCGCAACACGAGTACCAACAAATTCAGATTCAGCTGTGTCCAACCGGCAAA 960  
QY 324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343  
DB 961 CAAGCTGTTATCATCAGCTCTCAGACAGGATCACCGCGGTATTAACCGGCTGTCAAA 1020  
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363  
DB 1021 TCTGGAATTAAGCATTTTCAAAACTTATGATCAAAAGCATTCGCCAAACCATTTCTTAA 1080  
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383  
DB 1081 GGAATCTTAAGTTTTCCTTAAGGAGACTCAATGATTGCGAAGACTTCCCAAGCTC 1140  
QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403  
DB 1141 TCGAAGTCAATCCGCTCTTACCACTTAATGGTCAAGGTTGGGGTGTGATTT 1200  
QY 404 AlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 423  
DB 1201 CGCGGCCCTGCTCTCGTAAAGGATTAATGCAAAATCCACTCCGAGATGCAACAAAAC 1260  
QY 424 ValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443  
DB 1261 GTTCGTCATTTTCAGAAAGATCGGAAAACTGCAGGCTGCGGCTGATATTTCTATAG 1320  
QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463  
DB 1321 TTCATCTCAATTTTGGACACAGCAAGTAATAATTTGCTCAAAAACAAACAGCGAGTAT 1380  
QY 464 GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAlaIle 483  
DB 1381 GAATGACTCAAAAACCTACCAAGCTGGGCGCTCAATCTTAAAGCGTATGCGCAATC 1440  
QY 484 SerGlyAlaIleAlaGlyAla 490  
DB 1441 AGCGAGCCATTCGTCGCGCA 1461  
RESULT 2  
US-08-809-326A-17



Tue Jan 28 10:21:43 2003

us-09-889-314-2.rml

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 27, 2003, 14:11:52 ; Search time 51 Seconds  
(without alignments)  
2982.582 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 2369  
Sequence: 1 DTNMSISSSGPDNOKNIMS.....LKAYNAISGALGAHKTNNF 496

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ .p2n.model -DEV=nlh  
-O=/cgn2.1/USPTO.spool/US09889314/runat.24012003.144541.10130/app.query.fasta.1.647  
-DB=Issued Patents.NA -QFMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=45  
-MODE=LOCAL -OUTFORMAT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09889314.ecgn1.1.17 -gunat.24012003.144541.10130 -NCPU=6 -ICPU=3  
-NO\_XIPYX -NO\_MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents.NA.\*  
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2316	97.8	1464	4	US-08-809-326A-3
2	2316	97.8	1947	4	US-08-809-326A-17
3	1282	54.1	1568	4	US-08-809-326A-10
4	1265	53.4	5438	4	US-08-809-326A-25
5	1262	53.3	1048	4	US-08-809-326A-9
6	1245	52.6	777	4	US-08-809-326A-7
7	1245	52.6	813	4	US-08-809-326A-4
8	1245	52.6	1296	4	US-08-809-326A-18
9	220.5	9.3	5393	2	US-08-591-079-9
10	218.5	9.2	3622	2	US-08-591-079-7
11	213.5	9.0	32768	4	US-09-961-527-71
12	203	8.6	46819	4	US-09-453-702B-72

13	199	8.4	4483	4	US-08-961-527-363	Sequence 363, App
14	193	8.1	1702	4	US-08-961-527-367	Sequence 367, App
15	192	8.1	2550	4	US-08-961-527-364	Sequence 364, App
16	188	7.9	1436	4	US-08-961-527-365	Sequence 365, App
17	186	7.9	7766	4	US-09-125-619-3	Sequence 3, App11
18	185	7.8	1743	2	US-08-591-079-1	Sequence 1, App1
19	172.5	7.3	869	4	US-08-961-527-369	Sequence 369, App
20	170	7.2	43360	4	US-09-453-702B-206	Sequence 206, App
21	170	7.2	43325	4	US-09-453-702B-261	Sequence 261, App
22	167	7.0	3744	4	US-08-961-527-368	Sequence 368, App
23	166.5	7.0	941	4	US-08-961-527-370	Sequence 370, App
24	165.5	7.0	750	4	US-08-961-527-372	Sequence 372, App
25	158.5	6.7	807	4	US-08-961-527-373	Sequence 373, App
26	157	6.6	1068	4	US-08-961-527-373	Sequence 373, App
27	155.5	6.6	4411529	4	US-09-103-840A-1	Sequence 1, App11
28	155	6.5	4403765	4	US-09-103-840A-2	Sequence 2, App11
29	152.5	6.4	3106	2	US-08-591-079-5	Sequence 5, App11
30	152	6.4	6530	2	US-08-146-930-1	Sequence 1, App11
31	152	6.4	6530	3	US-08-458-240-1	Sequence 1, App11
32	152	6.4	6530	5	PCT-US93-03993-1	Sequence 1, App11
33	151	6.4	2835	4	US-09-134-001C-1515	Sequence 1515, App
34	151	6.4	6414	4	US-09-134-001C-1626	Sequence 1626, App
35	150.5	6.4	3561	4	US-09-134-001C-1685	Sequence 1685, App
36	150	6.3	5118	4	US-08-669-785-3	Sequence 3, App11
37	148	6.2	6530	2	US-08-146-930-1	Sequence 1, App11
38	148	6.2	6530	3	US-08-458-240-1	Sequence 1, App11
39	148	6.2	6530	5	PCT-US93-03993-1	Sequence 1, App11
40	147.5	6.2	690	4	US-08-961-527-377	Sequence 377, App
41	147.5	6.2	4649	6	5183745-1	Sequence 377, App
42	147.5	6.2	6441	4	US-08-669-785-1	Sequence 1, App11
43	147.5	6.2	6441	6	5183745-5	Sequence 1, App11
44	147.5	6.2	8078	4	US-09-702-251-3	Sequence 3, App11
45	146.5	6.2	3543	4	US-09-206-942-64	Sequence 64, App1

ALIGNMENTS

RESULT 1  
US-08-809-326A-3  
Sequence 3, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Iizutsu, Hiroshi  
APPLICANT: Obata, Kazuhiko  
APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE  
TITLE OF INVENTION: PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penile & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:



OY 294 AAAAGCAA 301  
|||||||  
DB 25 AAAAGCAA 32

## RESULT 42

protein W09G12.6 [imported] - Caenorhabditis elegans  
E88637  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: E88637  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: E88637  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-178 <STO>  
A:Cross-references: GB:chr\_IV; PIDN:AAC04448.1; PID:g2911883; GSPDB:GN00022; CESP:W09G12  
C:Genetics:  
A:Gene: W09G12.6  
A:Map position: 4

Query Match 1.6%; Score 8; DB 2; Length 178;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGGAGA 303  
|||||||  
DB 124 AAGGAGA 131

## RESULT 43

AB5217  
hypothetical protein AT4g19200 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: AB5217  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: AB5001; MID:20083488; PMID:10617198  
A:Accession: AB5217  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-179 <STO>  
A:Cross-references: GB:NC\_001268; MID:g7268715; PIDN:CAB78922.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g19200  
A:Map position: 4

Query Match 1.6%; Score 8; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAAGAAA 306  
|||||||  
DB 99 GAAGAAA 106

## RESULT 44

E70536  
hypothetical protein RV0804 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: E70536  
R:Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MID:9825987; PMID:9634230  
A:Accession: E70536  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-209 <COL>  
A:Cross-references: GB:Z95618; GB:AL123456; MID:g3261788; PIDN:CAB09107.1; PID:g31727  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV0804

Query Match 1.6%; Score 8; DB 2; Length 209;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 300 AAGAAAT 307  
|||||||  
DB 63 AAGAAAT 70

## RESULT 45

G97631  
hypothetical protein AGR\_C\_4120 [imported] - Agrobacterium tumefaciens (strain C58, C  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: G97631  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; PMID:11743194  
A:Accession: G97631  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-214 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK8008.1; PID:g15157423; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_4120  
A:Map position: circular chromosome

Query Match 1.6%; Score 8; DB 2; Length 214;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 284 AGAAGAA 291  
|||||||  
DB 165 AGAAGAA 172

Search completed: January 27, 2003, 16:37:07  
Job time : 53 secs

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Yamaharan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: H75262  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1145 <MUT>  
 A:Cross-references: GB:AE002082; GB:AE000513; NID:96460347; PIDN:AAF12073.1; PID:9646035  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2529  
 A:Map position: 1

Query Match 1.6%; Score 8; DB 2; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 284 GLAGLAA 291  
 DB 33 GAGLAA 40  
 |||||

RESULT 38  
 S43893  
 H+-exporting ATPase (EC 3.6.3.6) lipid-binding protein - *Neurospora crassa*  
 C:Species: *Neurospora crassa*  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: S43893  
 R:Sista, H.; Wechsler, M.A.; Bowman, B.J.  
 Mol. Gen. Genet. 243, 82-90, 1994  
 A:Title: The proteolipid subunit of the *Neurospora crassa* vacuolar ATPase: isolation of  
 A:Reference number: S43893; MUID:94247360; PMID:8190074  
 A:Accession: S43893  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1161 <STS>  
 A:Cross-references: EMBL:L07105; NID:g168929; PIDN:AAA19974.1; PID:g168930  
 A:Note: the authors translated the codon ATC for residue 63 as Tyr and GCT for residue 1  
 C:Genetics:  
 A:Introns: 5/2; 9/3; 53/1; 141/1  
 C:Superfamily: vacuolar H<sup>+</sup>-transporting ATPase 16k chain  
 C:Keywords: ATP; hydrolase

Query Match 1.6%; Score 8; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGLAA 289  
 DB 99 GAGLAA 106  
 |||||

RESULT 39  
 S32970  
 H+-exporting ATPase (EC 3.6.3.6) chain c-6, vacuolar - fission yeast (*Schizosaccharomyces*  
 N:Alternate names: vacuolar H<sup>+</sup>-transporting ATPase c-6  
 C:Species: *Schizosaccharomyces pombe*  
 C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 03-Jun-2002  
 C:Accession: S32970; F38033  
 R:Toyama, R.; Goldstein, D.J.; Schlegel, R.; Dhar, R.  
 Yeast 7, 989-991, 1991  
 A:Title: A genomic sequence of the *Schizosaccharomyces pombe* 16 kDa vacuolar H<sup>(+)</sup>-ATPase  
 A:Reference number: S32970; MUID:92206078; PMID:1839480  
 A:Accession: S32970  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1161 <TOY>  
 A:Cross-references: EMBL:X59947; NID:g5135; PIDN:CAA42572.1; PID:g5136  
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: Z21763

A:Accession: T38033  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1161 <MUR>  
 A:Cross-references: EMBL:Z98598; PIDN:CAB11240.1; GSPDB:GN00066; SPDB:SPAC13.14  
 A:Experimental source: strain 972h-; consold clB3  
 C:Genetics:  
 A:Gene: SPAC13.14  
 A:Map position: 1  
 A:Introns: 6/2; 52/2; 160/3  
 C:Superfamily: vacuolar H<sup>+</sup>-transporting ATPase 16k chain  
 C:Keywords: ATP; hydrogen ion transport; hydrolase; membrane-associated complex; tran

Query Match 1.6%; Score 8; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGLAA 289  
 DB 99 GAGLAA 106  
 |||||

RESULT 40  
 A40814  
 H+-exporting ATPase (EC 3.6.3.6) proteolipid chain, vacuolar - oat  
 C:Species: *Avena sativa* (oat)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: A40814  
 R:Lal, S.; Watson, J.C.; Hansen, J.N.; Size, H.  
 J. Biol. Chem. 266, 16078-16084, 1991  
 A:Title: Molecular cloning and sequencing of cDNAs encoding the proteolipid subunit o  
 A:Reference number: A40814; MUID:91340758; PMID:1831453  
 A:Accession: A40814  
 A:Molecule type: mRNA  
 A:Residues: 1165 <LAT>  
 A:Cross-references: GB:M73232; NID:g166548; PIDN:AAA32712.1; PID:g166549  
 C:Superfamily: vacuolar H<sup>+</sup>-transporting ATPase 16k chain  
 C:Keywords: ATP; hydrogen ion transport; hydrolase; transmembrane protein

Query Match 1.6%; Score 8; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGLAA 289  
 DB 103 GAGLAA 110  
 |||||

RESULT 41  
 G87307  
 hypothetical protein CC0472 [imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
 C:Accession: G87307  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: G87307  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1178 <STO>  
 A:Cross-references: GB:AE005673; NID:g13421647; PIDN:AAK22459.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC0472  
 C:Superfamily: ubiquinol-cytochrome c reductase iron-sulfur protein; Rieske [2Fe-2S]

Query Match 1.6%; Score 8; DB 2; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Cross-references: GB:X02632; NID:g5686; PIDN:CAA26479.1; PID:g5687  
 R:Ammons, R.; Puijms, W.; Moller, W.  
 FEBS Lett. 104, 85-89, 1979  
 A:Title: The primary structure of ribosomal protein eL12/eL12-P from *Artemia salina* 80 S  
 A:Reference number: A02775; MUID:80004136; PMID:477981  
 A:Accession: A02775  
 A:Molecule type: protein  
 A:Residues: 1-12, 'T', 14-18, 'TX', 21-33, 'X', 35-79, 'A', 81-111 <AMO>  
 A>Note: Ser-98 is partially phosphorylated; this results in two slightly different forms  
 C:Superfamily: rat acidic ribosomal protein P1  
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome  
 F:1-111/Product: ribosomal protein L12el #status experimental <MAT>  
 F:98/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 1.6%; Score 8; DB 1; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301  
 |||||  
 Db 72 AAAAGGAA 79

RESULT 33  
 S41866  
 acidic ribosomal protein P2 - fungus (*Cladosporium herbarum*)  
 C:Species: *Cladosporium herbarum*  
 C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 13-Aug-1999  
 C:Accession: S41866  
 R:Zhang, L.  
 submitted to the EMBL Data Library, January 1994  
 A:Description: A cDNA clone coding for a novel allergen, Cla hIII, of *Cladosporium herbarum*  
 A:Reference number: S41866  
 A:Accession: S41866  
 A:Molecule type: mRNA  
 A:Residues: 1-111 <ZHA>  
 A:Cross-references: EMBL:X77253; NID:g452605; PIDN:CAA54470.1; PID:g452606  
 C:Superfamily: rat acidic ribosomal protein P1  
 C:Keywords: protein biosynthesis; ribosome  
 F:1-111/Product: acidic ribosomal protein P2.e.B #status predicted <MAT>  
 F:70-89/Region: alanine-rich  
 F:99-111/Region: aspartic acid/glutamic acid-rich

Query Match 1.6%; Score 8; DB 2; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGA 300  
 |||||  
 Db 78 GAAAGGA 85

RESULT 34  
 R0072B  
 acidic ribosomal protein P2-B - *Trypanosoma cruzi*  
 C:Species: *Trypanosoma cruzi*  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 22-Jun-1999  
 C:Accession: S22950; S19948  
 R:Vaquer, M.P.; Schijman, A.G.; Panabera, A.; Levin, M.J.  
 Nucleic Acids Res. 20, 2893, 1992  
 A:Title: Acidic sequence of a cDNA encoding another *Trypanosoma cruzi* acidic ribosomal  
 A:Reference number: S22950; MUID:92310999; PMID:1614880  
 A:Accession: S22950  
 A:Molecule type: mRNA  
 A:Residues: 1-112 <VAZ>  
 A:Cross-references: EMBL:X65065; NID:g10633; PIDN:CAA46198.1; PID:g10634  
 A>Note: It is uncertain whether Met-1 or Met-3 is the initiator  
 C:Superfamily: rat acidic ribosomal protein P1  
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 1; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301  
 |||||  
 Db 81 AAAAGGAA 88

RESULT 35  
 G70354  
 hypothetical protein aq\_615 - *Aquifex aeolicus*  
 C:Species: *Aquifex aeolicus*  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
 C:Accession: G70354  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
 A:Reference number: A70300; MUID:98196666; PMID:9537320  
 A:Accession: G70354  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-132 <AOF>  
 A:Cross-references: GB:AE000698; NID:g2983224; PIDN:AAC06830.1; PID:g2983233; GB:AE00  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: aq\_615  
 C:Superfamily: *Aquifex aeolicus* hypothetical protein aq\_615

Query Match 1.6%; Score 8; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAA 296  
 |||||  
 Db 35 GAAVGAA 42

RESULT 36  
 A83149  
 hypothetical protein PA3971 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 01-Mar-2002  
 C:Accession: A83149  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: A83149  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-143 <STO>  
 A:Cross-references: GB:AE004815; GB:AE004091; NID:g9950159; PIDN:AA007358.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA3971  
 C:Superfamily: hypothetical protein b1396

Query Match 1.6%; Score 8; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 300 AAGAAAT 307  
 |||||  
 Db 65 AAGAAAT 72

RESULT 37  
 H75262  
 hypothetical protein - *Deinococcus radiodurans* (strain R1)  
 C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: H75262

OY 290 AAVGAAA 297  
 |||||||  
 DB 63 AAVGAAA 70

## RESULT 29

T52147  
 ribosomal protein rplal [imported] - yeast (candida albicans)

C:Species: Candida albicans  
 C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 03-Nov-2000

C:Accession: T52147  
 R:Taylor, K.; Harris, D.

submitted to the EMBL Data Library, November 1998  
 A:Reference number: Z25985

A:Accession: T52147  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-106 <TAV>

A:Cross-references: EMBL:AL033497; PIDN:CAA21967.1

A:Experimental source: strain 1161; cosmid CA49C10

C:Genetics:

A:Gene position: 1

A:Superfamily: rat acidic ribosomal protein P1

C:Keywords: protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGA 301  
 |||||||  
 DB 67 AAAAGGA 74

## RESULT 30

R6BY22  
 60s acidic ribosomal protein p2.2 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: acidic ribosomal protein 2; ribosomal protein SP-140c; ribosomal prot

C:Species: Schizosaccharomyces pombe  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 10-Dec-1999

C:Accession: B34715; A27304; S10054; T40800

R:Beltrame, M.; Bianchi, M.E.

Mol. Cell. Biol. 10, 2341-2348, 1990

A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: two e

A:Reference number: A34715; MUID:90220620; PMID:2325655

A:Accession: B34715  
 A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:M33138; NID:g173465; PIDN:AAA35335.1; PID:g173466

R:Beltrame, M.; Bianchi, M.E.

Nucleic Acids Res. 15, 9089, 1987

A:Title: Sequence of the cDNA for one acidic ribosomal protein of Schizosaccharomyces po

A:Reference number: A27304; MUID:88067727; PMID:3684587

A:Accession: A27304  
 A:Molecule type: mRNA

A:Residues: 1-110 <BEL2>

A:Cross-references: GB:Y00466; NID:g5064; PIDN:CAA6528.1; PID:g5065

R:Okada, E.; Higo, K.I.; Itoh, T.

Mol. Gen. Genet. 193, 519-524, 1983

A:Title: Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from Schizosaccha

A:Reference number: S07293; MUID:84038947; PMID:6355773

A:Accession: S10054  
 A:Molecule type: protein

A:Residues: 1-15, 'B', '17-21', 'B', '23', 'Z', '25-32', 'Z', '34', 'ZXZ', '38-39', 'Z', <OTA>

R:Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z21949  
 A:Accession: T40800

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-110 <BEC>

A:Cross-references: EMBL:AL032684; PIDN:CAA21791.1; GSPDB:GN00067; SPDB:SPBP87.06

A:Experimental source: strain 972h; clone p1 p887

C:Genetics:

A:Gene: rpa2

A:Map position: 2

C:Superfamily: rat acidic ribosomal protein P1

C:Keywords: phosphoprotein; protein biosynthesis; ribosome

F,1-110/Product: acidic ribosomal protein P2.2 #status experimental <MAT>

Query Match 1.6%; Score 8; DB 1; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGA 301  
 |||||||  
 DB 76 AAAAGGA 83

## RESULT 31

R6BY24  
 60s acidic ribosomal protein p2-beta - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: acidic ribosomal protein 4; ribosomal protein Y112el

C:Species: Schizosaccharomyces pombe  
 C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 10-Dec-1999

C:Accession: D34715; T39961

R:Beltrame, M.; Bianchi, M.E.

Mol. Cell. Biol. 10, 2341-2348, 1990

A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: tw

A:Reference number: A34715; MUID:90220620; PMID:2325655

A:Accession: D34715  
 A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: EMBL:M33142; NID:g173465; PIDN:AAA35337.1; PID:g173470

R:Xiang, Z.; Aves, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z21893  
 A:Accession: T39961

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-110 <LIA>

A:Cross-references: EMBL:AL035065; PIDN:CAA22631.1; GSPDB:GN00067; SPDB:SPBC2367.15c

A:Experimental source: strain 972h; cosmid c2367

C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 1; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGGAGA 303  
 |||||||  
 DB 75 AAGGAGA 82

## RESULT 32

R8S512  
 ribosomal protein L12el - brine shrimp

N:Alternate names: ribosomal protein E12

C:Species: Artemia salina (brine shrimp)

C:Date: 30-Nov-1979 #sequence\_revision 31-Mar-1992 #text\_change 22-Jun-1999

C:Accession: A25208; A02775

R:Maassen, J.A.; Schop, E.N.; Brands, J.H.G.M.; Van Hemert, F.J.; Lenstra, J.A.; Mol1

Eur. J. Biochem. 149, 609-616, 1985

A:Title: Molecular cloning and analysis of cDNA sequences for two ribosomal proteins

A:Reference number: A91146; MUID:85230659; PMID:3839187

A:Accession: A25208  
 A:Molecule type: mRNA

A:Residues: 1-108 <MAA>

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID:98295987; PMID:9634230

A: Accession: G70668

A: Status: nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-1616 <COL>

A: Cross-references: GB:Z83858; GB:AL123456; NID:q3261675; PIDN:CAB06103.1; PID:q1781167

A: Experimental source: strain H37Rv

R: Smith, D.R.; Robison, K.

Submitted to the EMBL Data Library, September 1994

A: Description: Mycobacterium tuberculosis cosmid tbcc2.

A: Reference number: S73053

A: Accession: S73075

A: Molecule type: DNA

A: Residues: 'MNGARMPVRCSTEL', 48-917, 'RSEARSRCGSGRAPAPPIGPPIGP', 'PMPPIIPVEPAGPVFN

A: Cross-references: EMBL:U00024; NID:q560506; PIDN:AAA50928.1; PID:q560507

C: Genetics:

A: Gene: pks1

C: Superfamily: Mycobacterium tuberculosis probable polyketide synthase pks1; acyl carrierology

C: Keywords: carrier protein

F: 84-362/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F: 1228-1410/Domain: short-chain alcohol dehydrogenase homology <SADH>

F: 1514-1585/Domain: acyl carrier protein homology <ACPI>

Query Match 1.8%; Score 9; DB 2; Length 1616;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGLAGLAG 289  
|||||  
Db 58 AGLAGLAG 66

RESULT 25

T26216

hypothetical protein W06A7.3c - Caenorhabditis elegans

C: Species: Caenorhabditis elegans

C: Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C: Accession: T26216

R: Ainscough, R.

Submitted to the EMBL Data Library, August 1996

A: Reference number: Z20173

A: Accession: T26216

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2484 <NLL>

A: Cross-references: EMBL:Z78066; PIDN:CAB51467.1; GSPDB:GN00023; CESP:W06A7.3c

A: Experimental source: clone W06A7

C: Genetics:

A: Gene: CESP:W06A7.3c

A: Map position: 5

A: Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2

Query Match 1.8%; Score 9; DB 2; Length 2484;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAGCAAA 297  
|||||  
Db 2207 GAAGCAAA 2215

RESULT 26

T26215

hypothetical protein W06A7.3a - Caenorhabditis elegans

C: Species: Caenorhabditis elegans

C: Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C: Accession: T26215

R: Ainscough, R.

Submitted to the EMBL Data Library, August 1996

A: Reference number: Z20173

A: Accession: T26215

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2607 <NLL>

A: Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a

A: Experimental source: clone W06A7

C: Genetics:

A: Gene: CESP:W06A7.3a

A: Map position: 5

A: Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2; 2586/2

Query Match 1.8%; Score 9; DB 2; Length 2607;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAGCAAA 297  
|||||  
Db 2207 GAAGCAAA 2215

RESULT 27

T31328

fibroin - Chinese oak silkworm

C: Species: Antheraea pernyi (Chinese oak silkworm)

C: Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999

C: Accession: T31328

R: Sezutsu, H.; Tamura, T.; Yukuhira, K.

Submitted to the EMBL Data Library, August 1998

A: Description: Characterization of the full length fibroin gene of a wild silkworm, A

A: Reference number: Z20995

A: Accession: T31328

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2639 <SE2>

A: Cross-references: EMBL:AF083334; NID:q3450882; PID:q3450883; PIDN:AC32606.1

C: Genetics:

A: Introns: 14/3

Query Match 1.8%; Score 9; DB 2; Length 2639;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 297 AGGAGAAA 305  
|||||  
Db 117 AGGAGAAA 125

RESULT 28

T35731

hypothetical protein SC7H1.36c SC7H1.36c - Streptomyces coelicolor (fragment)

C: Species: Streptomyces coelicolor

C: Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C: Accession: T35731

R: Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

Submitted to the EMBL Data Library, January 1998

A: Reference number: Z21548

A: Accession: T35731

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-81 <MRP>

A: Cross-references: EMBL:AL021411; PIDN:CA16223.1; GSPDB:GN00070; SCOPDB:SC7H1.36c

A: Experimental source: strain A3(2)

C: Genetics:

A: Gene: SCOPDB:SC7H1.36c

Query Match 1.6%; Score 8; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

F:28-860/Product: elastin #status predicted <MAT>  
F:850-855/Disulfide bonds: #status predicted

Query Match 1.8%; Score 9; DB 1; Length 860;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAA 306  
|||||  
Db 113 GGAAGAAA 121

## RESULT 20

140889  
sarcosine oxidase (EC 1.5.3.1) alpha chain [validated] - Corynebacterium sp.  
C:Species: Corynebacterium sp.  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 26-May-2000  
C:Accession: 140889

R:Chlumsky, L.J.; Zhang, L.; Jorns, M.S.

J. Biol. Chem. 270, 18252-18259, 1995

A:Title: Sequence analysis of sarcosine oxidase and nearby genes reveals homologues with  
A:Reference number: A57385; MUID:95355441; PMID:7543100

A:Accession: 140889

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-967 <RES>

A:Cross-references: EMBL:U23955; NID:927587; PIDN:AAC43461.1; PID:927591

C:Genetics:

A:Gene: soxa

C:Complex: heterotetramer; alpha (PIR:140889), beta (PIR:140887), gamma (PIR:140888), an

C:Function: EC 1.5.3.1 [validated, MUID:95355441]; catalyzes the oxidative demethylat

A:Description: EC 1.5.3.1 [validated, MUID:95355441]; FAD plus 1 mol FAD covalently attac

A:Note: enzyme contains 1 mol of noncovalently bound FAD plus 1 mol FAD covalently attac

C:Keywords: oxidoreductase

Query Match 1.8%; Score 9; DB 2; Length 967;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAAGA 297  
|||||  
Db 434 GAAVGAAGA 442

RESULT 21  
113807  
potassium channel protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: 113807

R:Goldstein, S.A.; Price, L.A.; Rosenthal, D.N.; Pausch, M.H.

Proc. Natl. Acad. Sci. U.S.A. 93, 13256-13261, 1996

A:Title: ORK1, a potassium-selective leak channel with two pore domains cloned from Dros

A:Reference number: 217770; MUID:97075152; PMID:8917578

A:Accession: 113807

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1001 <GOL>

A:Cross-references: EMBL:U55321; NID:93808067; PID:93808068; PIDN:AAC69250.1

C:Genetics:

A:Gene: ORK1

A:Cross-references: FlyBase:FBgn0017561

A:Map position: 1

Query Match 1.8%; Score 9; DB 2; Length 1001;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAG 302  
|||||  
Db 764 AAAAGGAG 772

## RESULT 22

B70985  
probable polyketide synthase - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 26-May-2000

C:Accession: B70985

R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70985

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1017 <COL>

A:Cross-references: GB:Z95617; GB:AL123456; NID:93242249; PIDN:CAB09100.1; PID:e31726

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: pks9

C:Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-

C:Keywords: carrier protein

F:24-392/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F:504-782/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F:864-942/Domain: acyl carrier protein homology <ACP2>

Query Match 1.8%; Score 9; DB 2; Length 1017;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGIAGIAG 289  
|||||  
Db 478 AGIAGIAG 486

RESULT 23  
T00117  
dve protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 17-Nov-2000

C:Accession: T00117

R:Nakagoshi, H.; Hoshi, M.; Nabeshima, Y.; Matsuzaki, F.

Genes Dev. 12, 2724-2734, 1998

A:Title: A novel homeobox gene mediates the DPP signal to establish functional specif

A:Reference number: Z14112; MUID:98402482; PMID:9732270

A:Accession: T00117

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1019 <NAK>

A:Cross-references: EMBL:AB010299; NID:93523072; PIDN:BA32660.1; PID:93523073

C:Genetics:

A:Gene: dve

A:Cross-references: FlyBase:FBgn0020307

A:Map position: 2

Query Match 1.8%; Score 9; DB 2; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAG 302  
|||||  
Db 477 AAAAGGAG 485

RESULT 24  
G70668  
polyketide synthase pks1 - Mycobacterium tuberculosis (strain H37Rv)

N:Alternate names: polyketide synthase pks002c

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: G70668; S73075

A:Experimental source: strain Nigg (MOPn)  
 C:Genetics:  
 A:Gene: TC0867  
 C:Superfamily: conserved hypothetical protein TC0867

Query Match  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 447 FWOQASKIA 455  
 Db 447 FWOQASKIA 455

## RESULT 17

EABO  
 N:Alternate names: splice form a - bovine  
 N:Contains: elastin precursor, splice form b; elastin precursor, splice form c  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 08-Jun-1989 #sequence\_revision 26-Jul-1996 #text\_change 22-Jun-1999  
 C:Accession: A31865; A26728; B26728; A22343; I45886  
 R:Yeh, H.; Anderson, N.; Ornstein-Goldstein, N.; Bashir, M.M.; Rosenbloom, J.C.; Abrams, Biochemistry 28, 2365-2370, 1989  
 A:Title: Structure of the bovine elastin gene and S1 nuclease analysis of alternative splicing  
 A:Reference number: A31865; MUID:89274159; PMID:2543440  
 A:Accession: A31865  
 A:Molecule type: DNA  
 A:Residues: 1-27 <YES>  
 A:Cross-references: GB:J02855; NID:g340504; PIDN:AAA30776.1; PID:g552339  
 J:Raju, K.; Anwar, R.A.  
 J. Biol. Chem. 262, 5755-5762, 1987  
 A:Title: Primary structures of bovine elastin a, b, and c deduced from the sequences of complementary DNAs  
 A:Reference number: A92640; MUID:87194772; PMID:3032943  
 A:Accession: A26728  
 A:Molecule type: mRNA  
 A:Residues: 1,'RS','4-11','E','13-636','V','638-747 <RAJ>  
 A:Cross-references: GB:J02717; NID:g163019; PIDN:AAA30503.1; PID:g163020  
 A:Accession: B26728  
 A:Molecule type: mRNA  
 A:Residues: 1,'RS','4-11','E','13-225,240-636','V','638-747 <RAJ>  
 A:Cross-references: GB:K03505; NID:g163025; PIDN:AAA30505.1; PID:g163026  
 A:Accession: C26728  
 A:Molecule type: mRNA  
 A:Residues: 1,'RS','4-11','E','13-225,260-636','V','638-747 <RAJ>  
 A:Cross-references: GB:K03506; NID:g163027; PIDN:AAA30506.1; PID:g163028  
 R:Cifila, G.; May, M.; Ornstein-Goldstein, N.; Indik, Z.; Morrow, S.; Yeh, H.S.; Rosenb Biochemistry 24, 3075-3080, 1985  
 A:Title: Structure of the 3' portion of the bovine elastin gene.  
 A:Reference number: A22343; MUID:85280426; PMID:2992576  
 A:Accession: A22343  
 A:Molecule type: DNA  
 A:Residues: 613-747 <CIC>  
 A:Cross-references: GB:M20415  
 R:Rosenbloom, J.  
 Lab. Invest. 51, 605-623, 1984  
 A:Title: Biology of disease: Elastin: Relation of protein and gene structure to disease.  
 A:Reference number: I45885; MUID:8505254; PMID:6150137  
 A:Accession: I45886  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 678-683,685-747 <ROS>  
 A:Cross-references: GB:M31898; NID:g163015; PIDN:AAA96417.1; PID:g163018  
 R:Brown, P.L.; Mechem, L.; Tisdale, C.; Mechem, R.P.  
 Biochem. Biophys. Res. Commun. 186, 549-555, 1992  
 A:Title: The cysteine residues in the carboxy terminal domain of tropoelastin form an intramolecular disulfide bond  
 A:Contents: annotation, disulfide bonds  
 C:Comment: The term tropoelastin refers to a soluble precursor form of the extracellular line oxidase activity.  
 C:Genetics:  
 A:Introns: 634/3; 653/3; 676/3; 689/3; 707/3; 716/3; 733/3  
 A:Note: the list of introns is incomplete

C:Superfamily: elastin  
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine  
 F:1-747/Product: elastin precursor, splice form a #status predicted <EPA>  
 F:1-225,260-747/Product: elastin precursor, splice form c #status predicted <EPC>  
 F:1-225,240-747/Product: elastin precursor, splice form b #status predicted <EPB>  
 F:1-26/DNA: signal sequence #status predicted <SIG>  
 F:27-747/Product: elastin #status predicted <MAT>  
 F:105,109,252,271,275,324,327,400,404,407,445,448,489,493,544,548,552,606,609,645,649  
 F:737-742/Disulfide bonds: #status experimental

Query Match  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 298 GGAAGAAA 306  
 Db 95 GGAAGAAA 103

## RESULT 18

tropoelastin - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 23-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 24-Oct-1997  
 C:Accession: S59623; A24758  
 R:Mauch, J.C.; Sandberg, L.B.; Roos, P.J.; Jimenez, F.; Christiano, A.M.; Deak, S.B.; Matrix Biol. 14, 635-641, 1994  
 A:Title: Extensive alternate exon usage at the 5' end of the sheep tropoelastin gene.  
 A:Reference number: S59623  
 A:Accession: S59623  
 A:Molecule type: mRNA  
 A:Status: preliminary; not compared with conceptual translation  
 A:Residues: 1-770 <MAU>  
 R:Yoon, K.; Davidson, J.M.; Boyd, C.; May, M.; Luvall, P.; Ornstein-Goldstein, N.; S Arch. Biochem. Biophys. 241, 684-691, 1985  
 A:Title: Analysis of the 3' region of the sheep elastin gene.  
 A:Reference number: A24758; MUID:85305763; PMID:3839997  
 A:Accession: A24758  
 A:Molecule type: mRNA  
 A:Residues: 655-669,671-716,732-770 <YOO>  
 C:Superfamily: elastin  
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine  
 F:760-765/Disulfide bonds: #status predicted

Query Match  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 298 GGAAGAAA 306  
 Db 97 GGAAGAAA 105

## RESULT 19

EAMS  
 N:Alternate names: tropoelastin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 18-Aug-1995 #sequence\_revision 16-Aug-1996 #text\_change 22-Jun-1999  
 R:Wyner, K.S.; Sechler, J.L.; Boyd, C.D.; Passmore, H.C.  
 Genomics 23, 125-131, 1994  
 A:Title: Use of an intron length polymorphism to localize the tropoelastin gene to mouse chromosome 10  
 A:Reference number: A55721; MUID:95130069; PMID:7829060  
 A:Accession: A55721  
 A:Molecule type: mRNA  
 A:Residues: 1-860 <WY>  
 A:Cross-references: GB:U08210; NID:g473273; PIDN:AAA80155.1; PID:g473274  
 C:Genetics:  
 A:Map position: 5  
 C:Superfamily: elastin  
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine  
 F:1-27/DNA: signal sequence #status predicted <SIG>

C:Superfamily: peripheral-type benzodiazepine receptor

Query Match 1.8%; Score 9; DB 2; Length 169;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTTVA 310  
|||||  
DB 117 GAAATTTVA 125

#### RESULT 12

T34870  
probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 08-Sep-2000

C:Accession: T34870

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21560

A:Accession: T34870

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-243 <SEP>

A:Cross-references: EMBL:ALJ09849; PIDN:CAB52862.1; GSPDB:GN00070; SCOEDB:SC3A3.05c

A:Experimental source: strain A3(2)

C:Genetics:

C:Superfamily: Streptomyces coelicolor probable membrane protein SC3A3.05c

OY 282 GIAGLAAGA 290  
|||||  
DB 63 GIAGLAAGA 71

#### RESULT 13

AE2972  
secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002

C:Accession: AE2972

R:Wood, D.W.; Sebhal, J.C.; Kaut, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AE2972

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-437 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAJ44195.1; PID:g17741773; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: prfB

A:Map position: linear chromosome

Query Match 1.8%; Score 9; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 283 LAGLAAGA 291  
|||||  
DB 19 LAGLAAGA 27

#### RESULT 14

F98310

rhizobium secretion protein rse (AF141932) [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: F98310

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: F98310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-474 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK90008.1; PID:g15159979; GSPDB:GN00170

A:Genetics:

A:Gene: AGR\_L\_2881

A:Map position: linear chromosome

Query Match 1.8%; Score 9; DB 2; Length 474;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 283 LAGLAAGA 291  
|||||  
DB 56 LAGLAAGA 64

#### RESULT 15

D71497

hypothetical protein CT578 - Chlamydia trachomatis (serotype D, strain UM3/Cx)

C:Species: Chlamydia trachomatis

C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 18-Aug-2000

C:Accession: D71497

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:9900809; PMID:9784136

A:Accession: D71497

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-487 <ARN>

A:Cross-references: GB:AE001328; GB:AE001273; NID:g3329015; PIDN:MAC68180.1; PID:g332

A:Experimental source: serotype D, strain UM-3/Cx

C:Genetics:

A:Gene: CT578

C:Superfamily: conserved hypothetical protein TC0867

Query Match 1.8%; Score 9; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 447 FMQOASKTA 455  
|||||  
DB 443 FMQOASKTA 451

#### RESULT 16

F81655

conserved hypothetical protein TC0867 [imported] - Chlamydia muridarum (strain Nig9)

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 18-Aug-2000

C:Accession: F81655

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: F81655

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <TEP>

A:Cross-references: GB:AE002353; GB:AE002160; NID:g7190891; PIDN:AAF39663.1; PID:g719



A:Residues: 1-111 <ACH>  
 A:Cross-references: EMBL:X78223; NID:g467626; PIDN:CAA55067.1; PID:g467627  
 C:Superfamily: rat acidic ribosomal protein p1

Query Match 1.8%; Score 9; DB 2; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301  
 DB 77 GAAAGGAA 85

#### RESULT 7

T49895  
 oleosin-like protein - Arabidopsis thaliana

N:Alternate names: protein T211.270

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 17-Nov-2000

C:Accession: T49895

R:Bevan, M.; Murphy, G.; Riddle, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le  
 submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T49895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <BEV>

A:Cross-references: EMBL:AL163912; GSPDB:GN00063; ATSP:T211.270

A:Experimental source: cultivar Columbia; BAC clone T211

C:Genetics:

A:Gene: ATSP:T211.270

A:Map position: 5

A:Introns: 106/2

C:Superfamily: oleosin

Query Match 1.8%; Score 9; DB 2; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301  
 DB 119 GAAAGGAA 127

#### RESULT 8

I38724  
 mitochondrial benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 29-Sep-1999

C:Accession: I38724

R:Yakovlev, A.G.; Ruffo, M.; Jurka, J.; Krueger, K.E.

Gene 155, 201-205, 1995

A:Title: Comparison of repetitive elements in the third intron of human and rodent mitoc

A:Reference number: I38724; MUID:95237610; PMID:7721091

A:Accession: I38724

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-169 <RES>

A:Cross-references: EMBL:U12421; NID:g529945; PIDN:AAA83252.1; PID:g529946

C:Genetics:

A:Gene: MBR

A:Introns: 61/2; 107/3

C:Superfamily: peripheral-type benzodiazepine receptor

C:Keywords: mitochondrion

Query Match 1.8%; Score 9; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310  
 DB 117 GAAATTVA 125

#### RESULT 9

S14257  
 benzodiazepine receptor, peripheral-type - human

C:Species: Homo sapiens (man)

C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 29-Sep-1999

C:Accession: S14257

R:Riond, J.; Mattei, M.G.; Koghad, M.; Dumont, X.; Guillemot, J.C.; Le Fur, G.; Caput

Eur. J. Biochem. 195, 305-311, 1991

A:Title: Molecular cloning and chromosomal localization of a human peripheral-type be

A:Reference number: S14257; MUID:91146565; PMID:1847678

A:Accession: S14257

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-169 <RTO>

A:Cross-references: GB:M36035; NID:g184333; PIDN:AAA03652.1; PID:g306883

C:Genetics:

A:Gene: GDB:B2RP

A:Cross-references: GDB:127347; OMIM:109610

A:Map position: 22q13.31-32qter

C:Superfamily: peripheral-type benzodiazepine receptor

C:Keywords: mitochondrion; transmembrane protein

Query Match 1.8%; Score 9; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310  
 DB 117 GAAATTVA 125

#### RESULT 10

I38105  
 peripheral benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 29-Sep-1999

C:Accession: I38105

R:Lin, D.; Chang, Y.J.; Strauss, J.F.; Miller, W.L.

Genomics 18, 643-650, 1993

A:Title: The human peripheral benzodiazepine receptor gene: cloning and characterizat

A:Reference number: A49361; MUID:94140364; PMID:8307574

A:Accession: I38105

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-169 <RES>

A:Cross-references: GB:L21954; NID:g483405; PIDN:AAA18228.1; PID:g488425

C:Genetics:

A:Introns: 61/2; 107/3

C:Superfamily: peripheral-type benzodiazepine receptor

Query Match 1.8%; Score 9; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310  
 DB 117 GAAATTVA 125

#### RESULT 11

JE0149

peripheral benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 10-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 21-Nov-1998

C:Accession: JE0149

R:Xu, G.H.; Hu, B.R.; Luo, C.Q.; Chen, R.Z.; Qiu, P.X.

Chinese Biochem. J. 13, 423-427, 1997

A:Title: Cloning of the cDNAs encoding human peripheral benzodiazepine receptor and i

A:Reference number: JE0149

A:Accession: JE0149

A:Molecule type: mRNA

A:Residues: 1-169 <XUA>

Db 481 SGATAGA 487

## RESULT 3

G86591  
 C7578 hypothetical protein [imported] - Chlamydomonas pneumoniae (strain J138)  
 C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: G86591  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:203030349; PMID:10871362  
 A:Accession: G86591  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-488 <STO>  
 A:Cross-references: GB:BA000008; NID:98979183; PIDN:BA99017.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: CPJ0809  
 C:Superfamily: conserved hypothetical protein TC0867

Query Match 98.2%; Score 487; DB 2; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNKNIMSOVLSTPGGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63  
 |||||||  
 Db 1 MSISSSGPDNKNIMSOVLSTPGGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 60  
 QY 64 ASGKDTSTTETAPQGVAAKESSESQKAGADTVGSAATATATATKIAMQTSI 123  
 |||||||  
 Db 61 ASGKDTSTTETAPQGVAAKESSESQKAGADTVGSAATATATKIAMQTSI 120  
 QY 124 EENAKSMESTLESLSAQAQKEVAAYVAALSGSSGSALEMPKPGVTPRSEYI 183  
 |||||||  
 Db 121 EENAKSMESTLESLSAQAQKEVAAYVAALSGSSGSALEMPKPGVTPRSEYI 180  
 QY 184 EIGLALAKAIQTLGEATKSLSNVASTQAQADQTNKLGLEKQAIKIDKREERYQEMKAAE 243  
 |||||||  
 Db 181 EIGLALAKAIQTLGEATKSLSNVASTQAQADQTNKLGLEKQAIKIDKREERYQEMKAAE 240  
 QY 244 QKSKDEGMTVNTVMIAVSAIYISYIAAFTCGAGIAGIAAAGAAAGGAAGA 303  
 |||||||  
 Db 241 QKSKDEGMTVNTVMIAVSAIYISYIAAFTCGAGIAGIAAAGAAAGGAAGA 300  
 QY 304 AAFTVATOTTVQAVVQAVQAVITAVROAITAIIKAAYVSGIKAFITLVKAIKAISK 363  
 |||||||  
 Db 301 AAFTVATOTTVQAVVQAVQAVITAVROAITAIIKAAYVSGIKAFITLVKAIKAISK 360  
 QY 364 GISKVFAKTQMTAKNFKPKLSKIVISLTSKWTGVGVVVAAPALCKGMOLSEMOON 423  
 |||||||  
 Db 361 GISKVFAKTQMTAKNFKPKLSKIVISLTSKWTGVGVVVAAPALCKGMOLSEMOON 420  
 QY 424 VAOPKKEVGLQAAADMIISFTQFMOQASKIASKOTGESNEMOKATFKIGAOLIKAYAAI 483  
 |||||||  
 Db 421 VAOPKKEVGLQAAADMIISFTQFMOQASKIASKOTGESNEMOKATFKIGAOLIKAYAAI 480  
 QY 484 SGATAGA 490  
 |||||||  
 Db 481 SGATAGA 487

## RESULT 4

A43742  
 female sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)  
 N:Alternate names: membrane protein fsh, 205K  
 N:Contents: female sterile homeotic protein, 110K  
 C:Species: Drosophila melanogaster  
 C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 20-Sep-1993  
 C:Accession: A43742; B43742

R:Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; David, I.B.  
 Dev. Biol. 134, 246-257, 1989  
 A:Title: The Drosophila fsh locus, a maternal effect homeotic gene, encodes apparent  
 A:Reference number: A43742; MUID:89276730; PMID:2567251  
 A:Accession: A43742  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2038 <HAV>

A:Cross-references: EMBL:M23221; NID:g157452; PIDN:AAA28540.1; PID:g157453  
 A:Accession: B43742  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1106 <HA2>  
 A:Cross-references: EMBL:M23222  
 C:Genetics:  
 A:Gene: fsh  
 A:Cross-references: FlyBase:FBgn0004656  
 C:Superfamily: unassigned bromodomain proteins; bromodomain homology  
 C:Keywords: alternative splicing; transmembrane protein  
 F:1-2038/Product: female sterile homeotic protein, 205K #status predicted <MA2>  
 F:1-1106/Product: female sterile homeotic protein, 110K #status predicted <MA2>  
 F:59-116/Domain: bromodomain homology <BR01>  
 F:503-560/Domain: bromodomain homology <BR02>

Query Match 2.2%; Score 11; DB 2; Length 2038;  
 Best Local Similarity 100.0%; Pred. No. 0.29;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAGAGAGAA 304  
 |||||||  
 Db 343 AAAGAGAGAA 353

## RESULT 5

130714  
 hypothetical protein 112L - Molluscum contagiosum virus 1  
 N:Alternate names: MC112L  
 C:Species: Molluscum contagiosum virus 1  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jul-2000  
 C:Accession: T30714  
 R:Senkevitch, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
 Science 273, 813-816, 1996  
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host  
 A:Reference number: 220876; MUID:96325459; PMID:8670425  
 A:Accession: T30714  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-128 <SEN>  
 A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:AAC55240.1; PID:g1492055  
 C:Genetics:  
 A:Note: MC112L

Query Match 2.0%; Score 10; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 AAGAAGGAAA 296  
 |||||||  
 Db 83 AAGAAGGAAA 92

## RESULT 6

S43115  
 acidic ribosomal protein P2 - fungus (Cladosporium herbarum)  
 C:Species: Cladosporium herbarum  
 C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 13-Aug-1999  
 C:Accession: S43115  
 R:Acheltz, G.; Oberkofler, H.; Simon, B.; Lechenauer, E.; Unger, A.; Kandler, D.; Prill  
 submitted to the EMBL Data Library, March 1994  
 A:Description: Molecular characterization of allergens of Cladosporium herbarum and A  
 A:Reference number: S43108  
 A:Accession: S43115  
 A:Molecule type: mRNA

```
979 6 1.2 155 2 F84367 hypothetical prote
980 6 1.2 155 2 T23925 hypothetical prote
981 6 1.2 156 2 S41771 glycine-rich RNA-b
982 6 1.2 156 2 T470712 hypothetical prote
983 6 1.2 156 2 T47739 hypothetical prote
984 6 1.2 156 2 T16531 hypothetical prote
985 6 1.2 156 2 G72591 hypothetical prote
986 6 1.2 157 2 S49448 oleosin - rape
987 6 1.2 157 2 T33892 hypothetical prote
988 6 1.2 157 2 S13803 homeotic protein A
989 6 1.2 157 2 F87574 CBS domain protein
990 6 1.2 157 2 B69538 conserved hypotet
991 6 1.2 157 2 AE2223 hypothetical prote
992 6 1.2 158 2 H75028 probable H+-transp
993 6 1.2 158 2 G75542 probable RNA methy
994 6 1.2 158 2 C32055 nifx protein - Azot
995 6 1.2 158 2 C86937 probable membrane
996 6 1.2 158 2 E87271 hypothetical prote
997 6 1.2 158 2 H85620 hypothetical prote
998 6 1.2 158 2 E97500 conserved hypotet
999 6 1.2 158 2 A12718 type 4 pilin - Elk
1000 6 1.2 159 2 A47699
```

## ALIGNMENTS

## RESULT 1

F72031

c5578 hypothetical protein - Chlamydomophila pneumoniae (strain CML029)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 18-Aug-2000

C:Accession: F72031

R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: F72031

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 &lt;ARN&gt;

A:Cross-references: GB:AE001662; GB:AE001363; NID:94377118; PIDN:AAD18947.1; PID:94377112

A:Experimental source: strain CML029

C:Genetics:

A:Gene: CPN0809

C:Superfamily: conserved hypothetical protein TC0867

```
Query Match          99.4%; Score 493; DB 2; Length 493;
Best local Similarity 100.0%; Pred. No. 0;
Matches 493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTAG 63
    |||||||
DB 1 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTAG 60
QY 64 ASGKDKTSSSTTKETAPQOGVAAGKSSSESQKAGADTVGSGAATTAASNTATIAAQTST 123
    |||||||
DB 61 ASGKDKTSSSTTKETAPQOGVAAGKSSSESQKAGADTVGSGAATTAASNTATIAAQTST 120
QY 124 EASKSMESTLESLOSLSAAQMEKEVAVVAAALSGKSSGSAKLETPELPKPGVTPRSEVI 183
    |||||||
DB 121 EASKSMESTLESLOSLSAAQMEKEVAVVAAALSGKSSGSAKLETPELPKPGVTPRSEVI 180
QY 184 EIGLALAKAIQTLGEATKSLSNVASTQAOADOTNKLGLEKQAIKIDKEREVEOEKMAAE 243
    |||||||
DB 181 EIGLALAKAIQTLGEATKSLSNVASTQAOADOTNKLGLEKQAIKIDKEREVEOEKMAAE 240
QY 244 QKSKDLEGMTDVTNTVMIAVSAITVISAALFTGAGIAGLAAGAAGAAAGAGA 303
    |||||||
DB 241 QKSKDLEGMTDVTNTVMIAVSAITVISAALFTGAGIAGLAAGAAGAAAGAGA 300
QY 304 AAATTVATQITVOAVVOAKQAVITAVROAITAIRAANKSGIKAPITKLVIKAAIAISK 363
    |||||||
```

```
DB 301 AAATTVATQITVOAVVOAKQAVITAVROAITAIRAANKSGIKAPITKLVIKAAIAISK 360
QY 364 GISKVFAKGTOMIAKNPKISKVSSLTSTKMTVGVGVVAAAPALGKIMOMOLSEMOON 423
DB 361 GISKVFAKGTOMIAKNPKISKVSSLTSTKMTVGVGVVAAAPALGKIMOMOLSEMOON 420
QY 424 VAOFQKEVGLQAADMIISMTQFMQOASKIASKQGESNEMTOKATKLGAGIILKAYAAI 483
    |||||||
DB 421 VAOFQKEVGLQAADMIISMTQFMQOASKIASKQGESNEMTOKATKLGAGIILKAYAAI 480
QY 484 SGAIAGAHKHTNMF 496
DB 481 SGAIAGAHKHTNMF 493
```

## RESULT 2

AB1507

conserved hypothetical protein CP1062 (imported) - Chlamydomophila pneumoniae (strain A

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 18-Aug-2000

C:Accession: AB1507

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: AB1507

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-488 &lt;REA&gt;

A:Cross-references: GB:AE002263; GB:AE002161; NID:97189971; PIDN:AAF38835.1; PID:9718

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CP1062

C:Superfamily: conserved hypothetical protein TC0867

```
Query Match          98.2%; Score 487; DB 2; Length 488;
Best local Similarity 100.0%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTAG 63
    |||||||
DB 1 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTAG 60
QY 64 ASGKDKTSSSTTKETAPQOGVAAGKSSSESQKAGADTVGSGAATTAASNTATIAAQTST 123
    |||||||
DB 61 ASGKDKTSSSTTKETAPQOGVAAGKSSSESQKAGADTVGSGAATTAASNTATIAAQTST 120
QY 124 EASKSMESTLESLOSLSAAQMEKEVAVVAAALSGKSSGSAKLETPELPKPGVTPRSEVI 183
    |||||||
DB 121 EASKSMESTLESLOSLSAAQMEKEVAVVAAALSGKSSGSAKLETPELPKPGVTPRSEVI 180
QY 184 EIGLALAKAIQTLGEATKSLSNVASTQAOADOTNKLGLEKQAIKIDKEREVEOEKMAAE 243
    |||||||
DB 181 EIGLALAKAIQTLGEATKSLSNVASTQAOADOTNKLGLEKQAIKIDKEREVEOEKMAAE 240
QY 244 QKSKDLEGMTDVTNTVMIAVSAITVISAALFTGAGIAGLAAGAAGAAAGAGA 303
    |||||||
DB 241 QKSKDLEGMTDVTNTVMIAVSAITVISAALFTGAGIAGLAAGAAGAAAGAGA 300
QY 304 AAATTVATQITVOAVVOAKQAVITAVROAITAIRAANKSGIKAPITKLVIKAAIAISK 363
    |||||||
DB 301 AAATTVATQITVOAVVOAKQAVITAVROAITAIRAANKSGIKAPITKLVIKAAIAISK 360
QY 364 GISKVFAKGTOMIAKNPKISKVSSLTSTKMTVGVGVVAAAPALGKIMOMOLSEMOON 423
    |||||||
DB 361 GISKVFAKGTOMIAKNPKISKVSSLTSTKMTVGVGVVAAAPALGKIMOMOLSEMOON 420
QY 424 VAOFQKEVGLQAADMIISMTQFMQOASKIASKQGESNEMTOKATKLGAGIILKAYAAI 483
    |||||||
DB 421 VAOFQKEVGLQAADMIISMTQFMQOASKIASKQGESNEMTOKATKLGAGIILKAYAAI 480
QY 484 SGAIAGAHKHTNMF 490
```

833	6	1.2	124	2	S31684	Ig heavy chain V r
834	6	1.2	124	2	T32532	hypothetical prote
835	6	1.2	124	2	E70754	hypothetical prote
836	6	1.2	124	2	H84436	hypothetical prote
837	6	1.2	124	2	B26872	hypothetical prote
838	6	1.2	124	2	T05508	hypothetical prote
839	6	1.2	125	2	D50866	probable thioredox
840	6	1.2	125	2	D97595	ribosomal protein
841	6	1.2	125	2	AC2817	50S ribosomal prot
842	6	1.2	126	2	F70964	hypothetical prote
843	6	1.2	126	2	D81448	aspartate 1-decarb
844	6	1.2	126	2	T02225	auxin-regulated pr
845	6	1.2	126	2	AC2901	succinate dehydrog
846	6	1.2	126	2	T16777	hypothetical prote
847	6	1.2	127	2	H82285	hypothetical prote
848	6	1.2	127	2	B70454	iron-sulfur cofact
849	6	1.2	128	2	T44282	iron-sulfur cofact
850	6	1.2	128	2	F83168	probable iron-bind
851	6	1.2	128	2	E70466	ribosomal protein
852	6	1.2	128	2	UC1273	ribosomal protein
853	6	1.2	128	2	T40344	hypothetical prote
854	6	1.2	128	2	D70742	phycoerythrin alph
855	6	1.2	128	2	S12900	LSU ribosomal prot
856	6	1.2	129	2	C69488	hypothetical prote
857	6	1.2	129	2	G32845	peph protein - Acl
858	6	1.2	129	2	T00003	ribosomal protein
859	6	1.2	129	2	A71530	ribosomal protein
860	6	1.2	130	2	H81684	probable glutaredo
861	6	1.2	130	2	G96802	probable prolinae-r
862	6	1.2	130	2	T49946	T cell receptor CK
863	6	1.2	131	2	S57871	ribosomal protein
864	6	1.2	131	2	C75536	hypothetical prote
865	6	1.2	131	2	A87355	hypothetical prote
866	6	1.2	131	2	A64649	hemoglobin alpha-t
867	6	1.2	132	1	HAFGH	T-cell receptor be
868	6	1.2	132	1	S30441	Ig heavy chain V r
869	6	1.2	132	2	A38911	hypothetical prote
870	6	1.2	132	2	D83003	T-cell receptor be
871	6	1.2	133	1	RMH07B	probable cdd prote
872	6	1.2	133	2	B70843	ribosomal protein
873	6	1.2	133	2	G42645	ribosomal protein
874	6	1.2	133	2	B81684	ribosomal protein
875	6	1.2	133	2	T11792	hypothetical prote
876	6	1.2	133	2	D87593	uncharacterized co
877	6	1.2	133	2	C97301	natriuretic peptid
878	6	1.2	134	1	AMHUB	gene A2 protein -
879	6	1.2	134	1	MXBPT5	hypothetical prote
880	6	1.2	134	2	B83943	hypothetical prote
881	6	1.2	134	2	S77785	hypothetical prote
882	6	1.2	134	2	A72730	hypothetical prote
883	6	1.2	135	2	S38386	T-cell receptor be
884	6	1.2	135	2	S31604	Ig heavy chain V r
885	6	1.2	135	2	G72726	probable ribosomal
886	6	1.2	135	2	A70659	hypothetical prote
887	6	1.2	135	2	AH2100	hypothetical prote
888	6	1.2	135	2	A42098	transcription fact
889	6	1.2	136	2	AH1006	50S ribosomal chai
890	6	1.2	136	2	A96905	uncharacterized sm
891	6	1.2	136	2	D82982	cytochrome c5 PA53
892	6	1.2	136	2	T47982	hypothetical prote
893	6	1.2	136	2	C95291	superoxide dismuta
894	6	1.2	136	2	S52360	superoxide dismuta
895	6	1.2	137	2	S52362	superoxide dismuta
896	6	1.2	137	2	S52365	DNA-binding protei
897	6	1.2	137	2	S24755	DNA-binding protei
898	6	1.2	137	2	S00903	DNA-binding protei
899	6	1.2	137	2	C90846	DNA-binding protei
900	6	1.2	137	2	B85704	hypothetical prote
901	6	1.2	137	2	A70653	chorion protein S1
902	6	1.2	138	2	A32998	envelope protein -
903	6	1.2	138	2	S24106	gene A2-A3 protein
904	6	1.2	138	2	B46348	
905	6	1.2	138	2		

hypothetical prote  
T-cell receptor al  
chorion protein S1  
probable membrane  
14 kDa cuticle str  
high mobility grou  
probable nutrient  
conserved hypothet  
lysozyme (EC 3.2.1  
pregnancy-specific  
probable lipoprote  
Ig heavy chain V r  
hemoglobin alpha-D  
ribosomal L1 prote  
chorion protein S1  
conserved hypothet  
hypothetical prote  
conserved hypothet  
hypothetical prote  
globin F-I - spoon  
hypothetical prote  
hypothetical prote  
holin - lactobacil  
hypothetical prote  
actinoxanthin prec  
hypothetical 15k p  
DNA transport mach  
hypothetical prote  
hypothetical prote  
probable membrane  
hypothetical prote  
hypothetical prote  
hypothetical prote  
conserved hypothet  
globin F-I - spoon  
hypothetical prote  
hypothetical prote  
holin - lactobacil  
hypothetical prote  
actinoxanthin prec  
hypothetical 15k p  
DNA transport mach  
hypothetical prote  
hypothetical prote  
probable membrane  
hypothetical prote  
hypothetical prote  
hypothetical prote  
conserved hypothet  
vanadium chloroper  
hypothetical prote  
globin IIA, extrac  
foii protein - Nei  
probable transcrip  
Ig heavy chain V r  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
cuticle protein 8  
homeobox A5 protei  
foii protein MMB6  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
exonuclease (EC 3.  
hypothetical prote  
protein 135, dauer  
15.9k membrane pro  
RNAse-pseudouridine  
tRNA-pseudouridine  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
RNA-binding protei  
probable membrane  
hypothetical prote  
hypothetical prote  
plastoquinone precu  
conserved hypothet  
hypothetical prote  
acetyl-CoA carboxy

687	6	1.2	80	2	S62132	H+-transporting tw	760	6	1.2	106	2	G97584	hypothetical prote
688	6	1.2	80	2	S72757	B1496_C2_181 prote	761	6	1.2	106	2	AC2694	hypothetical prote
689	6	1.2	81	2	G82561	hypothetical prote	762	6	1.2	106	2	H97475	hypothetical prote
690	6	1.2	82	2	J01181	outer envelope mem	763	6	1.2	108	2	S01368	hypothetical prote
691	6	1.2	82	2	E85891	conserved hypotnet	764	6	1.2	108	2	B72586	hypothetical prote
692	6	1.2	83	2	T23070	hypothetical prote	765	6	1.2	108	2	F82598	hypothetical prote
693	6	1.2	87	2	B82115	conserved hypotnet	766	6	1.2	109	2	T01654	ribosomal protein
694	6	1.2	87	2	E95371	conserved hypotnet	767	6	1.2	109	2	D72602	hypothetical prote
695	6	1.2	88	1	MMAD9	early E1b 9K prote	768	6	1.2	109	2	A72492	hypothetical prote
696	6	1.2	88	2	I40653	9K protein - Coxie	769	6	1.2	109	2	B86050	hypothetical prote
697	6	1.2	88	2	C82627	hypothetical prote	770	6	1.2	110	1	R6SSP2	acidic ribosomal p
698	6	1.2	89	2	H72091	ribosomal protein	771	6	1.2	110	1	R6BYP3	60s acidic ribosom
699	6	1.2	89	2	G86531	L28 ribosomal prot	772	6	1.2	110	1	S44110	Ig heavy chain V-D
700	6	1.2	89	2	T42967	hypothetical prote	773	6	1.2	110	2	S76061	hypothetical prote
701	6	1.2	89	2	G82606	hypothetical prote	774	6	1.2	110	2	G87283	conserved hypotnet
702	6	1.2	90	2	B81102	conserved hypotnet	775	6	1.2	111	2	G91043	detoX protein [imp
703	6	1.2	90	2	G81135	hypothetical prote	776	6	1.2	111	2	B85888	hypothetical prote
704	6	1.2	90	2	E70004	hypothetical prote	777	6	1.2	111	2	B85888	hypothetical prote
705	6	1.2	90	2	AB3261	transglycosylase-a	778	6	1.2	111	2	T28220	hypothetical prote
706	6	1.2	90	2	AB2192	hypothetical prote	779	6	1.2	111	2	B72722	hypothetical prote
707	6	1.2	90	2	AH1955	hypothetical prote	780	6	1.2	111	2	A87525	EF hand domain pro
708	6	1.2	91	1	S00060	phospholipid trans	781	6	1.2	111	2	E80111	hypothetical prote
709	6	1.2	91	2	AB1573	cardoxysome struct	782	6	1.2	112	2	S54179	acidic ribosomal p
710	6	1.2	91	2	AG1219	cardoxysome struct	783	6	1.2	112	2	JQ1063	glycine-rich prote
711	6	1.2	91	2	E71938	hypothetical prote	784	6	1.2	113	1	R6FFP2	acidic ribosomal p
712	6	1.2	91	2	AF2387	transglycosylase-a	785	6	1.2	113	1	R6DOP1	acidic ribosomal p
713	6	1.2	92	2	DA3674	US5 protein - huma	786	6	1.2	113	2	PQ0273	polyprotein - hepa
714	6	1.2	92	2	F71868	hypothetical prote	787	6	1.2	113	2	G70754	hypothetical prote
715	6	1.2	92	2	I64109	mercury reductase	788	6	1.2	114	2	S17387	T-cell receptor be
716	6	1.2	93	2	F97129	truncated flavodox	789	6	1.2	114	2	B34792	Ig heavy chain pre
717	6	1.2	93	2	A95377	hypothetical prote	790	6	1.2	114	2	T06885	ribosomal protein
718	6	1.2	93	2	AF3485	hypothetical prote	791	6	1.2	114	2	AG1240	ribosome-binding f
719	6	1.2	94	2	AD1222	probable carboxyso	792	6	1.2	114	2	DB4984	hypothetical prote
720	6	1.2	94	2	AG1575	probable carboxyso	793	6	1.2	114	2	S32479	hypothetical prote
721	6	1.2	94	2	A10759	probable propanedi	794	6	1.2	114	2	E72724	hypothetical prote
722	6	1.2	94	2	H82374	acetoactate synth	795	6	1.2	114	2	G69183	conserved hypotnet
723	6	1.2	94	2	A64863	hypothetical prote	796	6	1.2	115	2	AH3305	nodulation protein
724	6	1.2	95	2	G82686	phosphotransferase	797	6	1.2	115	2	S66924	probable membrane
725	6	1.2	95	2	AG1218	Salmonella typhimu	798	6	1.2	116	1	TPAF	insulin precursor
726	6	1.2	95	2	C75478	conserved hypotnet	799	6	1.2	116	2	S21979	Ig heavy chain V-g
727	6	1.2	95	2	S72597	hypothetical prote	800	6	1.2	116	2	S17080	Ig heavy chain V-g
728	6	1.2	95	2	AG2439	hypothetical prote	801	6	1.2	116	2	H71687	hypothetical prote
729	6	1.2	96	2	D70254	hypothetical prote	802	6	1.2	116	2	D95280	hypothetical prote
730	6	1.2	96	2	C82970	hypothetical prote	803	6	1.2	116	2	H95414	hypothetical prote
731	6	1.2	96	2	SA0081	GPI-anchored lymph	804	6	1.2	117	1	S60030	hemoglobin - Param
732	6	1.2	97	2	T30456	hypothetical prote	805	6	1.2	117	1	S60032	hemoglobin - Param
733	6	1.2	97	2	T49767	hypothetical prote	806	6	1.2	117	1	T02872	probable lipid tra
734	6	1.2	97	2	D95328	hypothetical prote	807	6	1.2	117	2	T09155	lipid transfer pro
735	6	1.2	98	2	S26940	Ig heavy chain V r	808	6	1.2	117	2	E34792	Ig heavy chain pre
736	6	1.2	98	2	S38266	Ig heavy chain V r	809	6	1.2	117	2	D90510	hypothetical prote
737	6	1.2	98	2	P96643	cuticle protein lm	810	6	1.2	118	2	G34792	Ig heavy chain pre
738	6	1.2	98	2	D75519	hypothetical prote	811	6	1.2	118	2	D69837	hypothetical prote
739	6	1.2	99	1	Q8BPB7	hypothetical prote	812	6	1.2	119	2	PL0089	Ig heavy chain V r
740	6	1.2	99	1	QOQYBW	host specificity p	813	6	1.2	119	2	AC1032	Ig heavy chain V r
741	6	1.2	99	2	S26800	hypothetical prote	814	6	1.2	119	2	T14932	probable exported
742	6	1.2	99	2	S12418	Ig heavy chain V r	815	6	1.2	119	2	A69133	hypothetical prote
743	6	1.2	99	2	S12412	Ig heavy chain V r	816	6	1.2	120	2	A34871	conserved hypotnet
744	6	1.2	99	2	B18185	hypothetical prote	817	6	1.2	120	2	PT0370	Ig kappa chain V r
745	6	1.2	100	2	S78056	hypothetical prote	818	6	1.2	120	2	AD3598	Ig mu chain precu
746	6	1.2	100	2	G70486	Ig heavy chain V r	819	6	1.2	121	2	T04093	hypothetical prote
747	6	1.2	100	2	A12465	integration host f	820	6	1.2	122	2	AF1497	phospholipid trans
748	6	1.2	100	2	T06305	hypothetical prote	821	6	1.2	122	2	AG1139	hypothetical prote
749	6	1.2	101	2	G71017	hypothetical prote	822	6	1.2	122	2	S10213	hypothetical prote
750	6	1.2	101	2	F87421	MADH-ubiquinone ox	823	6	1.2	122	2	T51036	hypothetical prote
751	6	1.2	101	2	S77790	hypothetical prote	824	6	1.2	122	2	F83167	hypothetical prote
752	6	1.2	102	2	S28361	hypothetical prote	825	6	1.2	122	2	C84320	hypothetical prote
753	6	1.2	102	2	S73799	probable lipoprote	826	6	1.2	122	2	E69980	hypothetical prote
754	6	1.2	102	2	T51524	hypothetical prote	827	6	1.2	123	2	F64083	ribosomal protein
755	6	1.2	103	2	F75334	probable dioxygena	828	6	1.2	123	2	H75430	conserved hypotnet
756	6	1.2	103	2	F70178	hypothetical prote	829	6	1.2	123	2	A75273	hypothetical prote
757	6	1.2	104	2	S38268	cuticle protein lm	830	6	1.2	123	2	T22970	hypothetical prote
758	6	1.2	105	2	DB9875	hypothetical prote	831	6	1.2	123	2	H70633	hypothetical prote
759	6	1.2	106	2	H32537	T-cell receptor be	832	6	1.2	124	1	R5RFRA	ribosomal protein

541	7	1.4	847	2	T52436	614	7	1.4	3716	2	E70969	probable PPE prote
542	7	1.4	855	2	G86189	615	7	1.4	4196	2	T43274	dynein heavy chain
543	7	1.4	856	2	C71956	616	7	1.4	4613	2	T117409	polyketide synthas
544	7	1.4	856	2	H64552	617	7	1.4	4644	1	A38905	dynein heavy chain
545	7	1.4	871	2	AB4177	618	7	1.4	5188	2	B85547	probable RTX fam1
546	7	1.4	886	2	AB4112	619	7	1.4	5291	2	E90596	hypothetical prote
547	7	1.4	892	2	U00424	620	7	1.4	6260	2	T30228	polyketide synthas
548	7	1.4	895	2	T32780	621	7	1.4	8363	2	T30225	polyketide synthas
549	7	1.4	898	2	T37037	622	7	1.4	10223	2	T30225	polyketide synthas
550	7	1.4	915	2	AC3477	623	7	1.4	26926	15	S59489	titin cardiac mus
551	7	1.4	923	2	T38398	624	6	1.2	15	24	S59489	steroid monooxygen
552	7	1.4	928	2	T47846	625	6	1.2	29	24	PC2305	gaegurin 6 - Korea
553	7	1.4	941	2	T37526	626	6	1.2	33	2	A06683	malate dehydrogena
554	7	1.4	959	2	S32016	627	6	1.2	34	2	P00408	RNA-directed RNA p
555	7	1.4	963	2	B83502	628	6	1.2	35	2	B42762	Ig heavy chain V r
556	7	1.4	966	2	G69189	629	6	1.2	39	2	S26939	C 3.4.25.1 proteas
557	7	1.4	969	2	AB3388	630	6	1.2	41	2	D55539	Ig heavy chain V r
558	7	1.4	974	1	A40213	631	6	1.2	42	2	S33485	hypothetical prote
559	7	1.4	987	2	I48373	632	6	1.2	44	2	S58438	hypothetical prote
560	7	1.4	992	1	GNMVR3	633	6	1.2	47	2	I40320	vAr-repressed prot
561	7	1.4	1001	2	G87385	634	6	1.2	47	2	PC4133	hypothetical 47 pr
562	7	1.4	1011	2	F70620	635	6	1.2	50	2	T14888	hypothetical prote
563	7	1.4	1012	2	T00958	636	6	1.2	51	2	T13196	hypothetical prote
564	7	1.4	1038	2	JC5497	637	6	1.2	55	2	E82522	hypothetical prote
565	7	1.4	1063	1	GNMV77	638	6	1.2	56	2	T44431	hypothetical prote
566	7	1.4	1066	2	G84746	639	6	1.2	57	2	T03654	ribosomal protein
567	7	1.4	1071	2	AB4785	640	6	1.2	59	2	AB6485	glyceralddehyde-3-P
568	7	1.4	1079	2	F84946	641	6	1.2	59	2	H24802	protein F2809.16 l
569	7	1.4	1079	2	B70807	642	6	1.2	60	1	N1EY15	citicic protein 64
570	7	1.4	1091	2	S33596	643	6	1.2	60	1	N1EY15	short neurotoxin 1
571	7	1.4	1104	2	A60999	644	6	1.2	60	1	N1EY1P	short neurotoxin 1
572	7	1.4	1104	2	T49735	645	6	1.2	60	1	N1AT1F	short neurotoxin 1
573	7	1.4	1112	2	T02848	646	6	1.2	60	2	A27580	short neurotoxin 1
574	7	1.4	1122	2	B26427	647	6	1.2	61	2	S38917	hypothetical prote
575	7	1.4	1127	2	A25018	648	6	1.2	62	2	G25866	short neurotoxin b
576	7	1.4	1176	2	C26427	649	6	1.2	62	2	S14925	cerebellar polyep
577	7	1.4	1189	1	GNMVM7	650	6	1.2	62	2	A33915	hypothetical prote
578	7	1.4	1199	2	A40670	651	6	1.2	64	2	S36975	hypothetical prote
579	7	1.4	1201	2	A57369	652	6	1.2	64	2	H98033	hypothetical prote
580	7	1.4	1205	2	AB2486	653	6	1.2	65	2	T06457	outer envelope mem
581	7	1.4	1218	2	AC3580	654	6	1.2	65	2	AD0239	probable membrane
582	7	1.4	1218	2	A26427	655	6	1.2	65	2	B84278	hypothetical prote
583	7	1.4	1218	2	A26588	656	6	1.2	66	2	G82667	hypothetical prote
584	7	1.4	1221	2	A44978	657	6	1.2	66	2	A13117	hypothetical prote
585	7	1.4	1249	2	S68431	658	6	1.2	66	2	XSSOP5	excisionase - Stre
586	7	1.4	1280	2	A39117	659	6	1.2	67	1	C81042	hypothetical prote
587	7	1.4	1292	2	T09229	660	6	1.2	67	2	E72780	probable ribosomal
588	7	1.4	1343	2	AF0611	661	6	1.2	68	2	S77757	phosphotransferase
589	7	1.4	1361	2	E70806	662	6	1.2	69	2	B69355	hypothetical prote
590	7	1.4	1405	2	H87230	663	6	1.2	69	2	AC2563	hypothetical prote
591	7	1.4	1406	2	T17429	664	6	1.2	69	2	AP3566	hypothetical prote
592	7	1.4	1440	2	T44872	665	6	1.2	70	2	C53578	brevinin-1Bc precu
593	7	1.4	1473	2	T13855	666	6	1.2	71	2	LMZMA	H+-transporting tw
594	7	1.4	1486	2	F86785	667	74	1	74	1	LMVFM	H+-transporting tw
595	7	1.4	1535	2	T52081	668	6	1.2	74	1	LMVFM	H+-transporting tw
596	7	1.4	1538	2	A46221	669	6	1.2	74	1	S70027	H+-transporting tw
597	7	1.4	1557	2	H70646	670	6	1.2	74	2	S17767	H+-transporting tw
598	7	1.4	1557	2	T13160	671	6	1.2	74	2	S25960	H+-transporting tw
599	7	1.4	1640	2	D86798	672	6	1.2	74	2	A33667	H+-transporting tw
600	7	1.4	1686	2	AB7692	673	6	1.2	74	2	AG5952	H+-transporting tw
601	7	1.4	1901	2	F70806	674	6	1.2	74	2	AG2533	H+-transporting tw
602	7	1.4	1938	1	A40997	675	6	1.2	75	2	I62129	hypothetical prote
603	7	1.4	1969	2	T08875	676	6	1.2	76	2	S58180	gene Mid1 protein
604	7	1.4	1978	2	T07081	677	6	1.2	76	2	AD3291	su11 protein - mou
605	7	1.4	2103	2	G86925	678	6	1.2	77	2	PC1017	hypothetical prote
606	7	1.4	2175	1	S03170	679	6	1.2	77	2	AC2805	somatotropin precu
607	7	1.4	2257	2	T09538	680	6	1.2	78	2	AB2548	hypothetical prote
608	7	1.4	2257	2	D86483	681	6	1.2	78	2	C95254	D-alanyl carrier p
609	7	1.4	2261	2	T07084	682	6	1.2	79	2	PN0641	hypothetical prote
610	7	1.4	2304	2	T07920	683	6	1.2	79	2	T42116	hypothetical prote
611	7	1.4	2359	2	EB6483	684	6	1.2	79	2	A95519	D-alanyl carrier p
612	7	1.4	3002	2	A47221	685	6	1.2	79	2		hypothetical prote
613	7	1.4	3436	2	S55659	686	6	1.2	79	2	AG3443	hypothetical prote

395	7	1.4	482	2	JC7583	468	7	1.4	591	2	T45713	hypothetical prote
396	7	1.4	483	2	AG1898	469	7	1.4	595	2	AP2632	conserved hypotnet
397	7	1.4	483	2	T21327	470	7	1.4	595	2	B97415	hypothetical prote
398	7	1.4	488	2	AB2984	471	7	1.4	595	2	B97531	branched-chain ami
399	7	1.4	489	2	C82282	472	7	1.4	595	2	AC2750	hypothetical prote
400	7	1.4	490	2	A96556	473	7	1.4	598	2	JN0573	polyadenylate-bind
401	7	1.4	496	2	H83592	474	7	1.4	598	2	T22610	hypothetical prote
402	7	1.4	502	2	A48679	475	7	1.4	599	2	T72682	trkA family protei
403	7	1.4	503	2	C71550	476	7	1.4	603	2	AB7342	conserved hypotnet
404	7	1.4	503	2	T30966	477	7	1.4	603	2	B98107	hypothetical prote
405	7	1.4	504	2	T34995	478	7	1.4	606	2	T08180	p20 protein, micr
406	7	1.4	507	2	T35677	479	7	1.4	610	2	C96732	hypothetical prote
407	7	1.4	509	2	D84349	480	7	1.4	616	2	C83404	conserved hypotnet
408	7	1.4	511	2	G87609	481	7	1.4	617	2	T01227	glutathione-regula
409	7	1.4	520	2	S72324	482	7	1.4	623	2	T52293	MYC-related DNA bi
410	7	1.4	524	2	E82309	483	7	1.4	627	2	A41609	dnak-type molecula
411	7	1.4	524	2	E87548	484	7	1.4	627	2	T10290	hypothetical prote
412	7	1.4	526	2	AB6440	485	7	1.4	631	2	T05103	hypothetical prote
413	7	1.4	528	2	C96564	486	7	1.4	633	2	T27499	hypothetical prote
414	7	1.4	528	2	D46449	487	7	1.4	637	2	S48921	hypothetical prote
415	7	1.4	530	2	B87369	488	7	1.4	638	1	XXAV	dihydrolipoamide S
416	7	1.4	532	2	A72694	489	7	1.4	640	2	T08179	LRG5 protein - chl
417	7	1.4	533	2	S18539	490	7	1.4	643	2	T43052	fatty acid transpo
418	7	1.4	535	2	S07521	491	7	1.4	648	2	B40727	S-M checkpoint con
419	7	1.4	536	2	T37544	492	7	1.4	649	2	S48379	glycerol-3-phospha
420	7	1.4	541	2	T23088	493	7	1.4	654	2	D87602	sensory box histid
421	7	1.4	541	2	AB7595	494	7	1.4	654	2	T34613	NADH2 dehydrogenas
422	7	1.4	541	2	S73383	495	7	1.4	657	2	D71351	probable primosoma
423	7	1.4	542	2	S42089	496	7	1.4	659	2	T02838	probable membrane
424	7	1.4	542	2	AB2356	497	7	1.4	666	2	D87243	conserved integral
425	7	1.4	542	2	F97369	498	7	1.4	671	2	C96534	probable Poly-A Bi
426	7	1.4	542	2	AP2587	499	7	1.4	677	2	G87586	cation transportin
427	7	1.4	545	2	B89823	500	7	1.4	682	2	AG3351	transporter BME107
428	7	1.4	547	2	C98267	501	7	1.4	693	2	G82618	plus biogenesis p
429	7	1.4	547	2	AE3017	502	7	1.4	696	2	A90959	probable tail prot
430	7	1.4	548	1	SYBYAL	503	7	1.4	696	2	C85807	hypothetical prote
431	7	1.4	549	1	D90477	504	7	1.4	696	2	A11029	hypothetical prote
432	7	1.4	553	2	JC7723	505	7	1.4	699	2	D70533	hypothetical prote
433	7	1.4	554	2	G75576	506	7	1.4	708	2	J01148	killer toxin KHS p
434	7	1.4	556	2	AD1770	507	7	1.4	733	2	D87112	primosomal protein
435	7	1.4	556	2	A11394	508	7	1.4	735	2	A59434	KIAA1501 protein
436	7	1.4	558	2	S52951	509	7	1.4	738	2	T00343	hypothetical prote
437	7	1.4	559	2	S52944	510	7	1.4	742	2	T35650	probable ATP-depen
438	7	1.4	559	2	S52944	511	7	1.4	749	2	B86606	primosomal protein
439	7	1.4	559	2	S52940	512	7	1.4	749	2	C72018	primosomal protein
440	7	1.4	559	2	S52937	513	7	1.4	750	2	T48804	hypothetical prote
441	7	1.4	559	2	S52950	514	7	1.4	752	2	T27276	hypothetical prote
442	7	1.4	559	2	S52935	515	7	1.4	752	2	AC3624	acriflavine resista
443	7	1.4	559	2	S52938	516	7	1.4	753	2	B71472	probable primosoma
444	7	1.4	559	2	S52956	517	7	1.4	753	2	D81734	primosomal protein
445	7	1.4	559	2	S52955	518	7	1.4	778	2	F70963	hypothetical glyci
446	7	1.4	559	2	S52954	519	7	1.4	778	2	D87965	protein V63D3A.6p
447	7	1.4	559	2	S52953	520	7	1.4	778	2	S56293	probable membrane
448	7	1.4	559	2	E84213	521	7	1.4	779	2	AB6862	primosomal protein
449	7	1.4	560	2	S52943	522	7	1.4	780	1	KIHUFM	6-phosphofructokin
450	7	1.4	560	2	H83206	523	7	1.4	780	1	KIRBF	6-phosphofructokin
451	7	1.4	564	2	F87305	524	7	1.4	780	2	S71429	phosphofructokin
452	7	1.4	566	2	T45162	525	7	1.4	781	2	A13014	methyl-accepting C
453	7	1.4	566	2	D70709	526	7	1.4	788	2	F98269	hypothetical prote
454	7	1.4	566	2	F71403	527	7	1.4	790	2	D83011	conserved hypotnet
455	7	1.4	570	1	C69985	528	7	1.4	795	2	B83608	hypothetical prote
456	7	1.4	570	2	A39597	529	7	1.4	797	2	AH1302	primosomal replica
457	7	1.4	570	2	S54134	530	7	1.4	797	2	AH1674	primosomal replica
458	7	1.4	571	2	S65060	531	7	1.4	798	2	C95202	primosomal protein
459	7	1.4	572	2	T00707	532	7	1.4	798	2	C98069	primosomal replica
460	7	1.4	573	2	C85433	533	7	1.4	802	2	G89893	primosomal P
461	7	1.4	574	2	S61234	534	7	1.4	803	2	B84931	DNA topoisomerase
462	7	1.4	575	2	B83586	535	7	1.4	804	2	E83963	primosomal replica
463	7	1.4	580	2	G83339	536	7	1.4	805	2	A69682	non-motile and pha
464	7	1.4	583	2	T10701	537	7	1.4	813	2	AC3304	hypothetical prote
465	7	1.4	584	2	G70804	538	7	1.4	814	2	T14608	hypothetical prote
466	7	1.4	588	2	G95242	539	7	1.4	828	2	T03544	hypothetical prote
467	7	1.4	588	2	T22688	540	7	1.4	839	2	F75518	hypothetical prote

249	7	1.4	316	2	F85826
250	7	1.4	317	2	T00146
251	7	1.4	317	2	AH0496
252	7	1.4	320	2	T42062
253	7	1.4	322	1	A53307
254	7	1.4	322	2	T36577
255	7	1.4	322	2	T40526
256	7	1.4	323	2	S20099
257	7	1.4	324	2	F90311
258	7	1.4	324	2	AH2844
259	7	1.4	324	2	T44374
260	7	1.4	328	2	AF2081
261	7	1.4	330	2	B87502
262	7	1.4	331	1	A30929
263	7	1.4	331	2	H64957
264	7	1.4	331	2	F90963
265	7	1.4	331	2	F85811
266	7	1.4	331	2	AC0752
267	7	1.4	331	2	F70820
268	7	1.4	331	2	AC3552
269	7	1.4	332	2	D70716
270	7	1.4	333	1	S44027
271	7	1.4	333	2	T41743
272	7	1.4	333	2	F48423
273	7	1.4	335	2	S53596
274	7	1.4	337	2	T06196
275	7	1.4	339	2	A83128
276	7	1.4	339	2	T34925
277	7	1.4	344	2	A70771
278	7	1.4	345	2	A64493
279	7	1.4	346	1	MUHGPD
280	7	1.4	346	2	AD3210
281	7	1.4	347	2	S44846
282	7	1.4	347	2	D91588
283	7	1.4	351	2	D64185
284	7	1.4	353	2	B70358
285	7	1.4	358	2	H97613
286	7	1.4	358	2	AE2836
287	7	1.4	359	2	I49341
288	7	1.4	359	2	T26813
289	7	1.4	361	2	H72587
290	7	1.4	362	2	S45094
291	7	1.4	364	2	C48376
292	7	1.4	365	2	T12040
293	7	1.4	365	2	F96986
294	7	1.4	366	2	T16119
295	7	1.4	366	2	T31820
296	7	1.4	366	2	A97451
297	7	1.4	366	2	AC2669
298	7	1.4	367	2	G71178
299	7	1.4	368	2	T50233
300	7	1.4	369	2	T42295
301	7	1.4	370	2	S60187
302	7	1.4	371	2	AC3440
303	7	1.4	375	2	B82933
304	7	1.4	376	2	A53384
305	7	1.4	377	2	I39582
306	7	1.4	378	2	S60758
307	7	1.4	378	2	JC4648
308	7	1.4	379	2	A48936
309	7	1.4	379	2	T35653
310	7	1.4	382	2	AE2754
311	7	1.4	382	2	C97535
312	7	1.4	382	2	C97535
313	7	1.4	383	2	A84552
314	7	1.4	384	2	S25771
315	7	1.4	388	2	F75162
316	7	1.4	388	2	AF1907
317	7	1.4	389	2	T43927
318	7	1.4	391	2	A40594
319	7	1.4	392	2	C72427
320	7	1.4	395	2	A46031
321	7	1.4	396	2	T45088

probable transcrip  
hypothetical prote  
probable ornithine  
thioredoxin-disulf  
thioredoxin-disulf  
thioredoxin reduct  
ade/atp translocas  
transforming prote  
hypothetical prote  
thioredoxin reduct  
hypothetical prote  
thioredoxin reduct  
conserved hypothet  
flagellar motor sw  
flagellar motor sw  
flagellar motor sw  
hypothetical prote  
flagellar motor sw  
hypothetical glycol  
high-affinity bran  
hypothetical prote  
thioredoxin-disulf  
thioredoxin-disulf  
homeotic protein, e  
chlorophyll a/b-bi  
hypothetical prote  
acetoac catabolism  
ABC transporter in  
probable hydrolase  
acyl carrier prote  
M4-(beta-N'-acetyl  
ABC transporter, s  
K06H7.6 protein -  
bifunctional carbo  
UDP-N-acetylglucos  
hydrogenase (EC 1.  
hypothetical prote  
lytic murein trans  
MIP-1 alpha recept  
hypothetical prote  
hypothetical prote  
cinamyl-alcohol d  
orf 5' to phbc -  
cysteine proteinas  
sugar ABC transpor  
hypothetical prote  
hypothetical prote  
GTP cyclonhydrolase  
hypothetical prote  
hypothetical prote  
homeobox protein -  
hypothetical prote  
peridinin-chloroph  
GTP cyclonhydrolase  
hypothetical prote  
nitrite reductase  
polysialic acid ca  
siala protein - Nel  
nitrite reductase  
nitrite reductase  
hypothetical prote  
probable acyl-CoA  
two component sens  
nitrogen regulatio  
probable thioredox  
gas1 protein - mou  
tryptophan synthas  
tryptophan synthas  
tryptophan synthas  
major flagellin pr  
pyruvate synthase  
immobilization sur  
pyruvate synthase

322	7	1.4	398	1	OZ2QAS
323	7	1.4	401	1	OZ2QAC
324	7	1.4	401	2	AD1264
325	7	1.4	401	2	AF1626
326	7	1.4	404	2	H87106
327	7	1.4	407	2	D84518
328	7	1.4	407	2	D84169
329	7	1.4	410	2	S77535
330	7	1.4	411	2	S53933
331	7	1.4	412	2	F93594
332	7	1.4	413	2	B83180
333	7	1.4	414	2	S52618
334	7	1.4	415	2	T36906
335	7	1.4	416	1	A41267
336	7	1.4	416	2	E83590
337	7	1.4	417	2	AD0907
338	7	1.4	419	2	E90677
339	7	1.4	419	2	A85528
340	7	1.4	419	2	S22651
341	7	1.4	420	2	S06579
342	7	1.4	420	2	T50585
343	7	1.4	420	2	E86162
344	7	1.4	422	2	D84403
345	7	1.4	422	2	H82418
346	7	1.4	423	2	E70931
347	7	1.4	425	2	A31026
348	7	1.4	426	1	Z4BPD
349	7	1.4	426	1	Z4BPF1
350	7	1.4	430	2	T35676
351	7	1.4	432	2	H75622
352	7	1.4	433	2	B84566
353	7	1.4	436	2	U01459
354	7	1.4	436	2	E70387
355	7	1.4	438	2	JC7758
356	7	1.4	439	2	B82114
357	7	1.4	440	2	G87354
358	7	1.4	441	2	A75361
359	7	1.4	445	2	T36069
360	7	1.4	446	2	AE0088
361	7	1.4	446	2	T45525
362	7	1.4	450	2	C87463
363	7	1.4	451	2	G82096
364	7	1.4	452	2	S76671
365	7	1.4	452	2	F66906
366	7	1.4	454	2	T16062
367	7	1.4	454	2	T26296
368	7	1.4	457	2	H70820
369	7	1.4	457	2	B64818
370	7	1.4	457	2	C90919
371	7	1.4	457	2	H85767
372	7	1.4	458	1	S77662
373	7	1.4	458	2	A30775
374	7	1.4	462	2	H82160
375	7	1.4	463	2	F70627
376	7	1.4	467	2	E70976
377	7	1.4	469	2	G72602
378	7	1.4	470	1	S30830
379	7	1.4	470	2	AF0771
380	7	1.4	474	2	AH3594
381	7	1.4	474	2	S52444
382	7	1.4	474	2	S41117
383	7	1.4	474	2	S16250
384	7	1.4	475	2	T16441
385	7	1.4	475	2	F85430
386	7	1.4	475	2	T03170
387	7	1.4	476	2	G64720
388	7	1.4	476	2	G90629
389	7	1.4	476	2	G85480
390	7	1.4	477	2	D75469
391	7	1.4	478	2	T12683
392	7	1.4	479	2	F87474
393	7	1.4	479	2	AB2035
394	7	1.4	481	2	T00863

circumsporozoite p  
circumsporozoite p  
ammonium transport  
ammonium transport  
probable integral  
probable serpin f1  
molybdenum cofacto  
hypothetical prote  
probable membrane  
probable transmemb  
probable MFS trans  
hypothetical prote  
probable integral  
transcription fact  
cytosine permease  
cytosine permease  
cytosine permease  
cytosine permease/  
cytosine transport  
finger protein (cl  
probable membrane  
hypothetical prote  
dihydroorotase f1m  
ATP-dependent RNA  
probable PPE prote  
probable membrane  
gene IV protein -  
gene IV protein -  
probable ABC-type  
integral membrane  
hypothetical prote  
B1 protein precur  
conserved hypothet  
OVCAL protein - mo  
flagellum-specific  
hypothetical prote  
hypothetical prote  
hypothetical prote  
H+-transporting tw  
WSC4 homolog f1m  
hypothetical prote  
sodium-dependent t  
hypothetical prote  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical glycol  
glucuronide permea  
glucuronide permea  
thioredoxin-disulf  
dihydroliipoamide a  
conserved hypothet  
probable narv prot  
hypothetical prote  
hypothetical prote  
fumarate reductase  
fumarate reductase  
flavellin p- Legion  
flagellin p- Legion  
phytoene dehydroge  
hypothetical prote  
serine C-palmitoyl  
hypothetical prote  
probable inner mem  
probable inner mem  
inner membrane tra  
probable regulator  
embryogenesis prot  
major facilitator  
phytoene desaturas  
hypothetical prote



103	7	1.4	85	2	C87715	hypothetical prote
104	7	1.4	92	1	B87456	ribosomal protein
105	7	1.4	93	1	A37086	small acid-soluble
106	7	1.4	98	2	H85391	probable thiorodox
107	7	1.4	100	2	G97672	hypothetical prote
108	7	1.4	100	2	A82897	conserved hypothet
109	7	1.4	106	1	R8BY28	acidic ribosomal p
110	7	1.4	107	2	S03703	12K protein - pea
111	7	1.4	110	1	R5BYA1	acidic ribosomal p
112	7	1.4	110	2	T37490	ribosomal protein
113	7	1.4	110	2	T01565	acidic ribosomal p
114	7	1.4	112	2	S35440	acidic ribosomal p
115	7	1.4	113	2	S43109	acidic ribosomal p
116	7	1.4	113	2	P00274	polypepten - hepa
117	7	1.4	118	2	D72524	hypothetical prote
118	7	1.4	118	2	F87447	hypothetical prote
119	7	1.4	120	2	AD1106	ribosomal protein
120	7	1.4	122	1	R5HG12	ribosomal protein
121	7	1.4	122	1	T30865	stbc protein - Sal
122	7	1.4	123	1	R5BS9	ribosomal protein
123	7	1.4	123	2	G97286	ribosomal protein
124	7	1.4	123	2	D64003	probable sodium-tr
125	7	1.4	124	2	I40348	ribosomal protein
126	7	1.4	124	2	I40350	ribosomal protein
127	7	1.4	124	2	AE3345	LSU ribosomal prot
128	7	1.4	126	2	G70685	hypothetical prote
129	7	1.4	127	2	H87310	ribosomal protein
130	7	1.4	130	2	D84353	hypothetical prote
131	7	1.4	130	2	E55042	hypothetical prote
132	7	1.4	131	2	S75262	sensory transducti
133	7	1.4	132	2	P00394	genome polyprotein
134	7	1.4	132	2	P00396	genome polyprotein
135	7	1.4	133	2	T31324	hypothetical prote
136	7	1.4	138	2	C87389	hypothetical prote
137	7	1.4	140	2	D97722	50S ribosomal prot
138	7	1.4	142	2	B32998	chorion protein S1
139	7	1.4	146	2	S15449	hemoglobin beta ch
140	7	1.4	148	2	S50653	hypothetical prote
141	7	1.4	149	2	H75462	hypothetical prote
142	7	1.4	152	2	T49126	hypothetical prote
143	7	1.4	154	2	S04440	fibinrial protein -
144	7	1.4	154	2	B31105	hypothetical prote
145	7	1.4	155	2	D87414	hypothetical prote
146	7	1.4	157	2	T02034	early light-induce
147	7	1.4	163	1	D84674	hypothetical prote
148	7	1.4	165	1	CU08	plastoeyanin precu
149	7	1.4	167	2	T04903	iron-sulfur cofact
150	7	1.4	168	2	T02036	early light-induce
151	7	1.4	168	2	G72679	hypothetical prote
152	7	1.4	171	2	S76927	hypothetical prote
153	7	1.4	171	2	B87628	conserved hypothet
154	7	1.4	173	2	AD0158	probable Mut famli
155	7	1.4	176	2	F71540	hypothetical prote
156	7	1.4	178	2	E77069	hypothetical prote
157	7	1.4	178	2	E82742	hypothetical prote
158	7	1.4	179	2	D85602	probable complemen
159	7	1.4	180	2	G82858	hypothetical prote
160	7	1.4	184	2	S12095	embryonic abundan
161	7	1.4	185	2	F84182	hypothetical prote
162	7	1.4	188	2	G96652	protein F23N19.15
163	7	1.4	190	2	C71421	hypothetical prote
164	7	1.4	192	2	H91103	probable transport
165	7	1.4	192	2	C85949	hypothetical prote
166	7	1.4	192	2	AC0875	probable membrane
167	7	1.4	194	2	S24297	chohan protein -
168	7	1.4	195	2	E87296	DNA-3-methyladeni
169	7	1.4	200	2	H83297	hypothetical prote
170	7	1.4	200	2	T04147	LEA protein - rice
171	7	1.4	205	2	F87623	hypothetical prote
172	7	1.4	206	2	AC0692	conserved hypothet
173	7	1.4	207	2	I38602	zinc finger protei
174	7	1.4	215	2	E86656	transcription regu
175	7	1.4	215	2	D84382	hypothetical prote
176	7	1.4	216	2	F69166	hypothetical prote
177	7	1.4	220	2	S30977	gene 32 protein -
178	7	1.4	225	2	A87624	hypothetical prote
179	7	1.4	225	2	T50707	utase accessory p
180	7	1.4	225	2	C70500	probable transcrip
181	7	1.4	227	1	SYAD82	hexon-associated p
182	7	1.4	227	1	SYAD85	hexon-associated p
183	7	1.4	229	1	CA3330	gene 7 protein - p
184	7	1.4	230	2	G89892	orotidine-5-phosph
185	7	1.4	230	2	S72714	Lepb1170_F2_64 pro
186	7	1.4	231	2	T21119	hypothetical prote
187	7	1.4	233	2	S39346	hexon-associated p
188	7	1.4	235	2	T19328	hypothetical prote
189	7	1.4	237	2	H75374	beta-phosphoglucom
190	7	1.4	243	2	S07997	lipoprotein trar p
191	7	1.4	243	2	A38901	lipoprotein trar p
192	7	1.4	243	2	G64843	probable peroxidase
193	7	1.4	244	2	S01757	lipoprotein trar p
194	7	1.4	244	2	F70834	hypothetical prote
195	7	1.4	244	2	AB3609	multidrug resistan
196	7	1.4	245	2	AE2195	hypothetical prote
197	7	1.4	248	2	H80792	hypothetical prote
198	7	1.4	248	2	G75214	methyltransferase
199	7	1.4	249	2	G81218	bioH protein, prob
200	7	1.4	251	2	AB1751	triose phosphate 1
201	7	1.4	251	2	A11381	triose phosphate 1
202	7	1.4	251	2	S20455	pgqc protein - Kle
203	7	1.4	259	2	AF3174	3'-oxoacyl-(acyl)-ca
204	7	1.4	260	2	T35293	probable cobalamin
205	7	1.4	262	2	S00275	tail fiber protein
206	7	1.4	266	2	P00393	genome polyprotein
207	7	1.4	267	2	B87578	hypothetical prote
208	7	1.4	267	2	E72690	probable spermidin
209	7	1.4	269	2	T24358	hypothetical prote
210	7	1.4	269	2	E97296	PUR operon repress
211	7	1.4	271	1	A34950	dihydropterate re
212	7	1.4	272	2	G96930	nitrogenase iron p
213	7	1.4	276	2	E72643	probable phycho-2
214	7	1.4	276	2	G90786	probable cytochrom
215	7	1.4	277	2	T44589	tyrosin biosynthes
216	7	1.4	278	2	T08647	hypothetical prote
217	7	1.4	279	2	T48106	hypothetical prote
218	7	1.4	279	2	G85646	hypothetical prote
219	7	1.4	281	2	AB1585	B. subtilis yxbp p
220	7	1.4	282	2	AC0226	probable membrane
221	7	1.4	283	1	A56649	dihydropterate sy
222	7	1.4	283	2	A75274	hypothetical prote
223	7	1.4	284	2	G72662	hypothetical prote
224	7	1.4	287	2	A54601	RNA-binding protei
225	7	1.4	289	2	T49154	DNA-binding WRKY-1
226	7	1.4	291	2	T34043	hypothetical prote
227	7	1.4	298	2	C96022	probable 3-hydroxy
228	7	1.4	299	2	T35765	hypothetical prote
229	7	1.4	300	2	S24057	ferritin 2 precurs
230	7	1.4	300	2	B75619	probable cobalamin
231	7	1.4	300	2	B84310	hypothetical prote
232	7	1.4	302	1	S44026	thioredoxin-disulf
233	7	1.4	304	2	AH0763	probable transcrip
234	7	1.4	306	2	T42011	ADP,ATP carrier pr
235	7	1.4	306	2	S08607	chorion protein s3
236	7	1.4	307	2	AB82029	ribokinase [import
237	7	1.4	309	2	A17809	hypothetical prote
238	7	1.4	312	2	C81795	probable hydrolase
239	7	1.4	312	2	F70044	hypothetical prote
240	7	1.4	313	2	A75468	4-hydroxybenzoate
241	7	1.4	313	2	E91006	penicillin-binding
242	7	1.4	313	2	F85850	penicillin-binding
243	7	1.4	313	2	E64981	penicillin-binding
244	7	1.4	314	2	C98193	sugar transport sy
245	7	1.4	314	2	AC3093	hypothetical prote
246	7	1.4	316	1	F64966	probable transcrip
247	7	1.4	316	2	B97622	thioredoxin reduct
248	7	1.4	316	2	A99981	probable transcrip



**THIS PAGE BLANK (USPTO)**

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID-7227;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OREGON R;  
 RA Nakagoshi H., Hoshi M., Nabeshima Y., Matsuzaki F.;  
 RT "A novel homeobox gene mediates the DPP signal to establish functional  
 RT specificity within target cells."  
 RL Genes Dev. 0:0-0(1998).  
 DR EMBL: AB010299; BAA32660.1; .  
 DR HSSP: P10037; 1A07.  
 DR FLYBase: FBgn0020307; dve.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox; 3.  
 DR SMART: SM00389; HOX; 2.  
 DR PROSITE: PS50071; HOMEBOX\_2; 2.  
 SO SEQUENCE 1019 AA; 111355 MW; 5D22A104A0D0F0CC CRC64;

Query Match 1.88; Score 9; DB 5; Length 1019;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAGGAAG 302  
 DB 477 AAAGGAAG 485

Search completed: January 27, 2003, 16:36:18  
 Job time : 56 secs

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smith D.R.;  
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Robinson K.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U00024; AAA50928.1; -;  
 DR HSSP; P25715; IMMA.  
 DR InterPro; IPR001227; Ac.transferase.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00698; Acyl\_transf.1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 FT NON\_TER 950 950  
 SQ SEQUENCE 950 AA; 97188 MW; 305507720FB50DD8 CRC64;

Query Match 1.8%; Score 9; DB 2; Length 950;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGLAGLAG 289  
 DB 27 AGLAGLAG 35

RESULT 43  
 0960M0 PRELIMINARY; PRT; 1013 AA.  
 ID 0960M0;  
 AC 0960M0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE LD45559P.  
 GN CG1869.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Fattah D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouenavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Geulker S.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY051988; AAK83412.1; -;  
 DR FlyBase; FBgn0035398; CG1869.  
 DR InterPro; IPR001579; Chitinase\_18/2.  
 DR InterPro; IPR002557; Chitin\_bind\_Pera.  
 DR InterPro; IPR001223; Glyco\_hydro\_18.  
 DR Pfam; PF01607; CBM\_14; 1.  
 DR Pfam; PF00704; Glyco\_hydro\_18; 2.  
 DR PRODOM; PD000471; Glyco\_hydro\_18; 2.  
 DR PROSITE; PS01095; CHITINASE\_18; UNKNOWN\_2.  
 SQ SEQUENCE 1013 AA; 113285 MW; 266981AD0685EC02 CRC64;

Query Match 1.8%; Score 9; DB 5; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301  
 DB 88 GAAAGGAA 96

RESULT 44  
 006586

ID 006586 PRELIMINARY; PRT; 1017 AA.  
 AC 006586;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical 106.4 kDa protein (polyketide synthase).  
 GN PKS9 OR RV1664 OR MTCY275.03 OR MT1704.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / OSHKOSH;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Bishai W., Ullrich T., Weidman J., Knout H., Gill J., Mikula A.,  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z95617; CAB09100.1; -;  
 DR EMBL; AE007033; AAK45971.1; -;  
 DR HSSP; P25715; IMMA.  
 DR TIGR; MT1704; -;  
 DR TuberculList; RV1664; -;  
 DR InterPro; IPR001227; Ac.transferase.  
 DR InterPro; IPR000794; ketoacyl-synt.  
 DR InterPro; IPR003880; Panine\_attach.  
 DR Pfam; PF00698; Acyl\_transf.1.  
 DR Pfam; PF00109; ketoacyl-synt.1.  
 DR Pfam; PF02801; ketoacyl-synt.C; 1.  
 DR Pfam; PF00550; pp-binding.1.  
 DR PROSITE; PS00075; ACP\_DOMAIN; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN\_1.  
 KW Hypothetical protein; Phosphopantetheine; Complete proteome.  
 SQ SEQUENCE 1017 AA; 106402 MW; B93A22D818AE0EC CRC64;

Query Match 1.8%; Score 9; DB 16; Length 1017;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGLAGLAG 289  
 DB 478 AGLAGLAG 486

RESULT 45  
 076143 PRELIMINARY; PRT; 1019 AA.  
 ID 076143;  
 AC 076143;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE DVE.  
 GN DVE OR CG5799.  
 OS Drosophila melanogaster (Fruit fly).

```

Query Match          1.8%; Score 9; DB 5; Length 696;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCAA 301
    |||||||
DB 88 GAAAGCAA 96

RESULT 39
O9LID0 PRELIMINARY: PRT; 746 AA.
AC O9LID0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Similarity to amine oxidase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT "AC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL; AP001307; BAB01917.1; -
DR InterPro: IPR002937; Amino-oxidase.
DR InterPro: IPR002025; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
SQ SEQUENCE 746 AA; 82350 MW; E0FD2C0119C4EF00 CRC64;

Query Match          1.8%; Score 9; DB 10; Length 746;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGTAA 288
    |||||||
DB 166 GAGLAGTAA 174

RESULT 40
O8W0N1 PRELIMINARY: PRT; 748 AA.
AC O8W0N1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 82.1 kDa protein.
CN P0681B1.1 OR B1085F01.16.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_Taxid=4530, 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0681B1.1";

```

```

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:B1085F01.1";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003022; BAB78635.1; -
DR EMBL; AP003330; BAB89980.1; -
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 748 AA; 82079 MW; E18542CC048E36F7 CRC64;

Query Match          1.8%; Score 9; DB 10; Length 748;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAGAGAAA 306
    |||||||
DB 23 GGAGAGAAA 31

RESULT 41
O9ESZ9 PRELIMINARY: PRT; 810 AA.
AC O9ESZ9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE FLN (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Green E.D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289665; AAF99336.1; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
FT NON_TER 1
FT 1
SQ SEQUENCE 810 AA; 67204 MW; 1ED61ED5AD4F40A2 CRC64;

Query Match          1.8%; Score 9; DB 11; Length 810;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAGAGAAA 306
    |||||||
DB 63 GGAGAGAAA 71

RESULT 42
O50470 PRELIMINARY: PRT; 950 AA.
AC O50470;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Pks002c (Fragment).
OC Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

```

Query Match 1.8%; Score 9; DB 4; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288  
 |||||  
 DB 31 GAGLAGLAA 39

## RESULT 36

OY 099K82 PRELIMINARY; PRT; 555 AA.  
 AC 099K82;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Similar to hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 RN NCBI\_TaxID=10090;  
 RP [1]  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC004831; AAH04831.1; -  
 DR InterPro: IPR000759; Adrnck\_reductase.  
 DR InterPro: IPR002837; Amino\_oxidase.  
 DR InterPro: IPR002025; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PRINTS: PR00419; ADXRDTASE.  
 DR Hypothetical protein.  
 SO SEQUENCE 555 AA; 61852 MW; A297E9DBD094EA74 CRC64;

Query Match 1.8%; Score 9; DB 11; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288  
 |||||  
 DB 31 GAGLAGLAA 39

## RESULT 37

OY 091TN6 PRELIMINARY; PRT; 575 AA.

AC 091TN6;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE T52.

OS Tupaiia herpesvirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae.  
 RN NCBI\_TaxID=10397;  
 RP [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=2;  
 RA MEDLINE=21211637; PubMed=11312357;  
 RA Bahr U.; Darel G.;  
 RT "Analysis and Characterization of the Complete Genome of Tupaiia (Tree  
 RT Shrew) Herpesvirus."  
 RL J. Virol. 75:4854-4870(2001).  
 RN [2]  
 RN SEQUENCE FROM N.A.

RC STRAIN=2;  
 RA Darel G.; Bahr U.;  
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF281817; AAK57101.1; -  
 DR InterPro: IPR002597; Herpes\_env.  
 DR Pfam: PF01673; Herpes\_env; 1.  
 SO SEQUENCE 575 AA; 63743 MW; BE62EEF35166825D CRC64;

Query Match 1.8%; Score 9; DB 12; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAGGAAAT 307  
 |||||  
 DB 106 GAGGAAAT 114

## RESULT 38

OY 09VZV2 PRELIMINARY; PRT; 696 AA.  
 AC 09VZV2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE CG1869 protein.  
 GN CG1869.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID=7227;  
 RP [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;  
 DR MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu L.E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idagwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003477; AAF47714.1; -  
 DR FLYbase: FBgn0035398; CG1869.  
 DR InterPro: IPR001579; Chitinase\_18/2.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 2.  
 DR ProDom: PD000471; Glyco\_hydro\_18; 2.  
 DR PROSITE: PS01095; CHITINASE\_18; UNKNOWN\_1.  
 SO SEQUENCE 696 AA; 77366 MW; 6877585837A19CF6 CRC64;

DR EMBL: AE002353; AAF39663.1; -.  
 DR TIGR: TC0867; -.  
 DR PRINTS: PRO1608; BACINVASINC.  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 491 AA; 50762 MW; 440F298930D93FB2 CRC64;

Query Match 1.8%; Score 9; DB 16; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 FWOQASKIA 455  
 DB 447 FWOQASKIA 455

## RESULT 32

Q96LC4 PRELIMINARY; PRT; 502 AA.  
 AC Q96LC4; 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Polyamine oxidase isoform-2.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang Y., Devereux W., Stewart T.M., Casero R.A. Jr.;  
 RT "Human polyamine oxidase, isoform-2."  
 RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY033890; AAK55764.1; -.  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PRINTS: PRO0757; AMINEOXDASEF.  
 SQ SEQUENCE 502 AA; 56156 MW; A6CA22481CFEED4 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288  
 DB 31 GAGLAGLAA 39

## RESULT 33

Q9BW38 PRELIMINARY; PRT; 502 AA.  
 AC Q9BW38; 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Similar to hypothetical protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RT Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC000669; AA00669.1; -.  
 DR InterPro: IPR000759; Adrxn\_reductase.  
 DR InterPro: IPR001613; Aminoacid fl.  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PRINTS: PRO0419; ADXRDTASE.  
 KW Hypothetical protein.

SQ SEQUENCE 502 AA; 56091 MW; 45921536D1B92403 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288  
 DB 31 GAGLAGLAA 39

## RESULT 34

Q9NMW0 PRELIMINARY; PRT; 555 AA.  
 AC Q9NMW0; 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CDNA FLJ20746 fis, clone HEP06040.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kawakami T., Noguichi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,  
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;  
 RT "NEBO human cDNA sequencing project."  
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK000753; BAA91360.1; -.  
 DR InterPro: IPR000759; Adrxn\_reductase.  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PRINTS: PRO0419; ADXRDTASE.  
 SQ SEQUENCE 555 AA; 61819 MW; BDBEA65ECB9FA5BF CRC64;

Query Match 1.8%; Score 9; DB 4; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288  
 DB 31 GAGLAGLAA 39

## RESULT 35

Q96GT3 PRELIMINARY; PRT; 555 AA.  
 AC Q96GT3; 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Polyamine oxidase isoform-1.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21347206; PubMed=11454677;  
 RA Wang Y., Devereux W., Mosier P.M., Stewart T.M., Hacker A.,  
 RA Casero R.A. Jr.;  
 RT "Cloning and characterization of a human polyamine oxidase that is  
 inducible by polyamine analogue exposure."  
 RT Cancer Res. 61:5370-5373(2001).  
 RL EMBL: AY033889; AAK55763.1; -.  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PRINTS: PRO0757; AMINEOXDASEF.  
 SQ SEQUENCE 555 AA; 61872 MW; F0087D3629013C51 CRC64;



```

RESULT 28
O9SF45 PRELIMINARY: PRT: 477 AA.
ID 09SF45
AC 09SF45:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative oxidoreductase.
GN F11F8.15 OR AT3G09580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA:
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maltl R., Romling C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 3 BAC F11F8 genomic sequence.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M.; Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Kosena E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narasaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At3g09580 (GI:15232718) ";
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC016661; AAF23289.1; -
DR EMBL: AY064000; AAL36356.1; -
DR InterPro: IPR001613; Aminoacid-fl.
DR InterPro: IPR002937; Amino.oxidase.
DR Pfam: PF01593; Amino.oxidase; 1.
DR PRINTS: PR00757; AMINOXDASEF.
SQ SEQUENCE 477 AA: 52258 MW: C332782957852BE9 CRC64;
Query Match 1.8%; Score 9; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 280 GAGLAGLAA 288
Db 53 GAGLAGLAA 61

```

```

DR EMBL: AE001328; AAC68180.1; -.
KN Hypothetical protein: Complete proteome.
SQ SEQUENCE 487 AA: 50217 MW: DAD6038EC1DA03D CRC64;
Query Match 1.8%; Score 9; DB 16; Length 487;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 447 FWOQASKIA 455
Db 443 FWOQASKIA 451

```

---

```

RESULT 30
O9FTX7 PRELIMINARY: PRT: 491 AA.
ID 09FTX7
AC 09FTX7:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GH25390P.
GN ORK1 OR CG1615.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY058455; AAL13684.1; -
DR Flybase: FBgn0017561; ORK1.
SQ SEQUENCE 491 AA: 50962 MW: 282FBE74A3E90506 CRC64;
Query Match 1.8%; Score 9; DB 5; Length 491;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 294 AAAAGGAAG 302
Db 254 AAAAGGAAG 262

```

---

```

RESULT 31
O9PUG2 PRELIMINARY: PRT: 491 AA.
ID 09PUG2
AC 09PUG2:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein TC0867.
GN TC0867.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).

```

```
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
DR PRINTS: PR00419; ADXRPTASE
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR NON_TER 285 285
SQ SEQUENCE 285 AA; 31164 MW; 3E7622A42FCCE35A CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 2; Length 285;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
Db 7 GAGLAGLAA 15

RESULT 25
Q9XM22 PRELIMINARY; PRT; 307 AA.
AC Q9XM22;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ADP/ATP translocase.
GN AAC.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OG Mitochondrion.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Harmyeh S.H., Komuniecki R.;
RT "Identification of a cDNA coding for ADP/ATP translocase in the
RT parasitic nematode Ascaris suum.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF130365; AAB50505.1;
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARTRER.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
KW Mitochondrion.
SQ SEQUENCE 307 AA; 33575 MW; BD575E3AAF21F3F5-CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 8; Length 307;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AAGGAAGAA 304
Db 132 AAGGAAGAA 140

RESULT 26
Q9NPY2 PRELIMINARY; PRT; 308 AA.
AC Q9NPY2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DJ779E11.1.1 (novel flavin containing amine oxidase (isoform 1)).
GN DJ779E11.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL121675; CAC00598.1; -.
```

```
DR InterPro: IPR000759; Adnrx_reductase.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR000960; Flav_cont_moxgn.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
DR PRINTS: PR00419; ADXRPTASE.
DR PRINTS: PR00370; FMOXYGENASE.
SQ SEQUENCE 308 AA; 33734 MW; FCD4BF469B23D84F CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 4; Length 308;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
Db 31 GAGLAGLAA 39

RESULT 27
Q8UAJ1 PRELIMINARY; PRT; 474 AA.
AC Q8UAJ1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Secretion protein, HlyD family.
GN PRSE OR ATG3382 OR AGR L.2881.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Moks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,
RA Kutayvin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nestor E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gallung S., Miller N., Blanchard M., Mullin L.,
RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Marzelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL: AE009268; AAL44195.1; ALT_INIT.
DR EMBL: AE008344; AAK90008.1; -.
KW Complete proteome.
SQ SEQUENCE 474 AA; 52661 MW; 8BD17141B786EAC7 CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 16; Length 474;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 IAGLAAGAA 291
Db 56 IAGLAAGAA 64
```

RA Kormanec J., Homeroova D., Sevcikova B., Novakova R., Blaskovic D.;  
 RT "Large subunit of glutamate synthase of Streptomyces coelicolor  
 A3(2)";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF162938; AAD49328.1; -.  
 DR InterPro: IPR002839; DUF125.  
 DR Pfam: PF01988; DUF125; 1.  
 FT NON\_TER  
 SQ SEQUENCE 239 AA; 24777 MW; 5559563D7E5E962C CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 9; DB 2; Length 239;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 282 GLAAGAA 290  
 Db 59 GLAAGAA 67

RESULT 22  
 O9S2Y8 PRELIMINARY; PRT; 243 AA.  
 AC O9S2Y8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative membrane protein.  
 GN SC02027 OR SC3A3.05C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapalpe D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Latke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL109849; CAB52862.1; -.  
 DR InterPro: IPR002839; DUF125.  
 DR Pfam: PF01988; DUF125; 1.  
 SQ SEQUENCE 243 AA; 25205 MW; B3AF946838905DE1 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 9; DB 16; Length 243;  
 Matches 9; Conservative 100.0%; Pred. No. 6.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 282 GLAAGAA 290  
 Db 63 GLAAGAA 71

RESULT 23  
 O9K3G4 PRELIMINARY; PRT; 247 AA.  
 AC O9K3G4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical protein SC01277.  
 GN SC01277 OR 2SCG18.24.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Latke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL390188; CAB9157.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 247 AA; 22950 MW; EE32ABD76146881B CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 9; DB 16; Length 247;  
 Matches 9; Conservative 100.0%; Pred. No. 7;

Oy 283 LAGLAAGAA 291  
 Db 96 LAGLAAGAA 104

RESULT 24  
 O9RCT2 PRELIMINARY; PRT; 285 AA.  
 AC O9RCT2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Zea carotene desaturase (Fragment).  
 GN CRR0.  
 OS Synecchococcus leopoliensis.  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecchococcus.  
 OX NCBI\_TaxID=32047;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SAUG 1402-1;  
 RX MEDLINE=20465103; PubMed=11007968;  
 RA Miller B., Heuser T., Zimmer W.;  
 RT "Functional involvement of a deoxy-D-xylose 5-phosphate  
 reductoisomerase gene harboring locus of Synecchococcus leopoliensis in  
 isoprenoid biosynthesis";  
 RL FEBS Lett. 481:221-226(2000).  
 DR EMBL: AJ250721; CAB65434.1; -.  
 DR InterPro: IPR000759; Adnrx\_reductase.  
 DR InterPro: IPR001064; Crystallin.

```

DF 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Peripheral benzodiazepine receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=94140364; PubMed=8307574;
RA Lin D., Chang Y.J., Straus J.F., Miller W.L.;
RT "The human peripheral benzodiazepine receptor gene: cloning and
RT characterization of alternative splicing in normal tissues and in a
RT patient with congenital lipoid adrenal hyperplasia.";
RL Genomics 18:643-650(1993).
DR EMBL; L21954; AAA18228.1; -.
DR EMBL; L21952; AAA18228.1; JOINED.
DR EMBL; L21953; AAA18228.1; JOINED.
DR InterPro: IPR004307; Tspo_MBR.
DR Pfam: PF03073; Tspo_MBR; 1.
KW Receptor.
SQ SEQUENCE 169 AA; 18483 MW; C353A23AFADF327E CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 169;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310
Db 117 GAAATTTVA 125
|||||
RESULT 18
ID Q96TF6 PRELIMINARY; PRT; 169 AA.
AC Q96TF6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Df326114.1 (benzodiazepine receptor (peripheral)) (PBR, PKBS,
DE mitochondrial benzodiazepine, MBR) (isoform 1)).
GN BZRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Badbage A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z82214; CAB55884.1; -.
DR InterPro: IPR004307; Tspo_MBR.
DR Pfam: PF03073; Tspo_MBR; 1.
SQ SEQUENCE 169 AA; 18828 MW; 6DD741BF99A8B9D9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 169;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310
Db 117 GAAATTTVA 125
|||||
RESULT 19
ID Q9URS5 PRELIMINARY; PRT; 178 AA.
AC Q9URS5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE CRN1 homologue (Fragment).

```

```

GN CRN1.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2359/152;
RX MEDLINE=20136097; PubMed=10669871;
RA Bao W.G., Huo K.K., Li Y.Y., Fukuhara H.;
RT "Protein disulfide isomerase genes of Kluyveromyces lactis.";
RL Yeast 16:329-341(2000).
DR EMBL; AJ243960; CAB51777.1; -.
FT NON_TER
SQ SEQUENCE 178 AA; 18698 MW; C3D8E079CE461A CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 3; Length 178;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 GGAGAGAAA 306
Db 106 GGAGAGAAA 114
|||||
RESULT 20
ID Q96LC3 PRELIMINARY; PRT; 190 AA.
AC Q96LC3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Polyamine oxidase isoform-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Y., Devereux W., Stewart T.M., Hacker A., Casero R.A. Jr.;
RT "Human polyamine oxidase, isoform-3.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033891; AAK55765.1; -.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
SQ SEQUENCE 190 AA; 20616 MW; DE0614286373A56A CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 190;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
Db 31 GAGLAGLAA 39
|||||
RESULT 21
ID Q9S440 PRELIMINARY; PRT; 239 AA.
AC Q9S440;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE H3U (Fragment).
GN H3U.
OS Streptomyces coelicolor A3(2).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=100226;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M145;

```

ID 094F07 PRELIMINARY; PRT; 147 AA.  
 AC 094F07;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Glycine-rich protein GRP20.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21324873; PubMed=11431566;  
 RA Mayfield J.A., Flebig A., Johnstone S.E., Preuss D.;  
 RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";  
 RL Science 292:2482-2485(2001).  
 DR EMBL; AF362475; AAK83832.1; -  
 DR InterPro: IPR000136; Oleosin.  
 DR Pfam: PF01277; Oleosin; 1.  
 SO SEQUENCE 147 AA; 14878 MW; C3FEE84BC4F5C0F0 CRC64;

Query Match 1.8%; Score 9; DB 10; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 4.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301  
 DB 119 GAAAGGAA 127

RESULT 14  
 ID 094F02 PRELIMINARY; PRT; 149 AA.  
 AC 094F02;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Glycine-rich protein GRP20.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21324873; PubMed=11431566;  
 RA Mayfield J.A., Flebig A., Johnstone S.E., Preuss D.;  
 RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";  
 RL Science 292:2482-2485(2001).  
 DR EMBL; AF362476; AAK83838.1; -  
 DR InterPro: IPR000136; Oleosin.  
 DR Pfam: PF01277; Oleosin; 1.  
 SO SEQUENCE 149 AA; 15091 MW; FE5B139BAAD84E8 CRC64;

Query Match 1.8%; Score 9; DB 10; Length 149;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301  
 DB 119 GAAAGGAA 127

RESULT 15  
 ID 09LY07 PRELIMINARY; PRT; 153 AA.  
 AC 09LY07;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Oleosin-like protein (glycine-rich protein GRP20).  
 GN T211\_270.

OS Arabidopsis thaliana (Mouse-ear cross).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Beyer M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,  
 RA Rudd S., Lemcke K., Mayer K.F.X.;  
 RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21324873; PubMed=11431566;  
 RA Mayfield J.A., Flebig A., Johnstone S.E., Preuss D.;  
 RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";  
 RL Science 292:2482-2485(2001).  
 DR EMBL; AL163912; CAB87945.1; -  
 DR EMBL; AF362477; AAK83839.1; -  
 DR InterPro: IPR000136; Oleosin.  
 DR Pfam: PF01277; Oleosin; 1.  
 SO SEQUENCE 153 AA; 15458 MW; 8D04492C2F5E839B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 4.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301  
 DB 119 GAAAGGAA 127

RESULT 16  
 ID 094F02 PRELIMINARY; PRT; 153 AA.  
 AC 094F02;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Glycine-rich protein GRP20.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21324873; PubMed=11431566;  
 RA Mayfield J.A., Flebig A., Johnstone S.E., Preuss D.;  
 RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";  
 RL Science 292:2482-2485(2001).  
 DR EMBL; AF362474; AAK83826.1; -  
 DR InterPro: IPR000136; Oleosin.  
 DR Pfam: PF01277; Oleosin; 1.  
 SO SEQUENCE 153 AA; 15486 MW; 8D04492C3DD5B39B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 4.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301  
 DB 119 GAAAGGAA 127

RESULT 17  
 ID 013850 PRELIMINARY; PRT; 169 AA.  
 AC 013850;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)

Query Match 1.8%; Score 9; DB 3; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAG 302  
 |||||  
 DB 67 AAAAGGAG 75

## RESULT 9

O8TFM9 PRELIMINARY; PRT; 109 AA.  
 AC O8TFM9;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
 DE 60S acidic ribosomal protein P2.  
 OS Fusarium culmorum.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreales; mitosporic Hypocreales; Fusarium.  
 OX NCBI\_TaxID=5516;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hoff M., Krall M., Vieths S., Haustein D.;  
 RT "IGE-binding acidic ribosomal protein P2 is a potential allergen in  
 RT Fusarium culmorum.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY077706; AAT79930.1;  
 KM Ribosomal protein.  
 SQ SEQUENCE 109 AA; 11036 MW; D20E40DB073BDF77 CRC64;

Query Match 1.8%; Score 9; DB 3; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301  
 |||||  
 DB 76 GAAAGGAA 84

## RESULT 10

O9C3Z5 PRELIMINARY; PRT; 111 AA.  
 AC O9C3Z5;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
 DT 01-OCT-2001 (Tremblrel. 18, last annotation update)  
 DE Ribosomal protein P2.  
 OS Podospora anserina.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Lasiosphaeriaceae; Podospora.  
 OX NCBI\_TaxID=5145;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vierny C., Folichon M., Silar P.;  
 RT "Ribosomal proteins of Podospora anserina";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF331715; AKK11263.1;  
 DR InterPro: IPR001813; 60S-ribosomal.  
 DR Pfam: PF00428; 60S-ribosomal.P2.  
 DR PRINTS: PR00456; RIBOSOMALP2.  
 SQ SEQUENCE 111 AA; 10975 MW; 94D7379F9E73C414 CRC64;

Query Match 1.8%; Score 9; DB 3; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 3.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301  
 |||||  
 DB 72 GAAAGGAA 80

## RESULT 11

O8V7C9 PRELIMINARY; PRT; 128 AA.  
 AC O8V7C9;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
 DE ORF2.  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2184401; PubMed=11855633;  
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,  
 RA Okamoto H.;  
 RT "Analysis of the complete genomes of thirteen TT virus variants  
 RT classifiable into the fourth and fifth genetic groups, isolated from  
 RT viremic infants.";  
 RL Arch. Virol. 147:21-41(2002).  
 DR EMBL: AB064624; BAB79391.1;  
 DR InterPro: IPR004118; TT\_ORF2.  
 DR Pfam: PF02957; TT\_ORF2.1.  
 SQ SEQUENCE 128 AA; 13219 MW; 114D2C5DF694C37 CRC64;

Query Match 1.8%; Score 9; DB 12; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAGGAAA 306  
 |||||  
 DB 97 GGAGGAAA 105

## RESULT 12

O76068 PRELIMINARY; PRT; 143 AA.  
 AC O76068;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE Peripheral-type benzodiazepine receptor (Fragrant).  
 GN PBR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BREAST;  
 RA Hardwick M.J., Papadopoulos V.;  
 RT "Peripheral-type benzodiazepine Receptor (PBR) in Human Breast  
 RT Cancer.";  
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF075590; AAC31173.1;  
 DR EMBL: AF075589; AAC31172.1;  
 DR InterPro: IPR004307; TSPC\_MBR.  
 DR Pfam: PF03073; TSPC\_MBR.1.  
 FT NON\_TER  
 SQ SEQUENCE 143 AA; 16107 MW; DEAF74A8415C65A9 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTTA 310  
 |||||  
 DB 91 GAAATTTA 99

## RESULT 13

O94F07

RA Moss B.;  
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U60315; AAC55240.1; -;  
 DR InterPro: IPR004249; NPH3.  
 DR Pfam: PF03000; NPH3; 1.  
 SQ SEQUENCE 128 AA; 12728 MW; 22334EDB0B724607 CRC64;

Query Match 2.0%; Score 10; DB 12; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 0.46;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAGVAAA 296  
 |||||  
 DB 83 AAGAAGVAAA 92

## RESULT 5

ID 036896 PRELIMINARY; PRT; 137 AA.  
 AC 036896;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE GAG polyprotein (Fragment).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97445059; PubMed=9300048;  
 RA Leigh Brown A.J., Lobidel D., Wade C.M., Rebus S., Phillips N.,  
 RA Brettle R.P., France A.J., Leen C.S., McMenamin J., McMillan A.,  
 RA Maw R.D., Mulcahy F., Robertson J.R., Sankar K.N., Scott G., Wyld R.,  
 RA "The molecular epidemiology of human immunodeficiency virus type 1 in  
 RT six cities in Britain and Ireland."  
 RL Virology 235:166-177(1997).  
 DR EMBL: AF014293; AAC58374.1; -;  
 DR InterPro: IPR000071; Retrovir\_p17.  
 DR Pfam: PF00540; Gag\_p17; 1.  
 KW AIDS; Core protein; Polyprotein.  
 FT NON\_TER 1  
 FT NON\_TER 137  
 SQ SEQUENCE 137 AA; 14958 MW; 409CDC38A4F61522 CRC64;

Query Match 2.0%; Score 10; DB 15; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 0.49;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAGVAAA 296  
 |||||  
 DB 96 AAGAAGVAAA 105

## RESULT 6

ID 09F065 PRELIMINARY; PRT; 487 AA.  
 AC 09F065;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative non-phototropic hypocotyl 3.  
 GN P0013F10.9.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0013F10.";

RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AP002523; BAB17063.1; -;  
 DR InterPro: IPR004249; NPH3.  
 DR Pfam: PF03000; NPH3; 1.  
 SQ SEQUENCE 487 AA; 50508 MW; 84B57B75E64FC11 CRC64;

Query Match 2.0%; Score 10; DB 10; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAAT 307  
 |||||  
 DB 468 GGAAGAAAT 477

## RESULT 7

ID 08S076 PRELIMINARY; PRT; 103 AA.  
 AC 08S076;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE P0470A12.43 protein.  
 GN P0470A12.43.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0470A12.";  
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AP003436; BAB90315.1; -;  
 SQ SEQUENCE 103 AA; 10426 MW; 4A1EA91B1477B35B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAAG 302  
 |||||  
 DB 17 AAAAGGAAG 25

## RESULT 8

ID 09HF07 PRELIMINARY; PRT; 106 AA.  
 AC 09HF07;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE 60S acidic ribosomal protein type P1-A.  
 GN P1A.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCM 8215;  
 RA Abramczyk D., Tchorzewski M., Grankowski N.;  
 RT "Cloning, expression and purification of the acidic ribosomal protein  
 RT from Candida albicans.";  
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF317659; AAG33240.1; -;  
 DR InterPro: IPR001813; 60S\_ribosomal.  
 DR Pfam: PF00428; 60S\_ribosomal; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 106 AA; 10996 MW; BB03BB1B0CE9547E CRC64;





```
966 7 1.4 547 16 Q8U913 O8U913 agrobacteri
967 7 1.4 548 10 Q8W378 O8W378 oryza sativ
968 7 1.4 549 12 Q918H6 Q918H6 baboon cyto
969 7 1.4 549 17 Q97UN3 Q97UN3 sulfolobus
970 7 1.4 551 10 Q9FV45 Q9FV45 tagetes ere
971 7 1.4 551 16 Q92SW7 Q92SW7 rhizobium m
972 7 1.4 552 10 Q94KE8 Q94KE8 citrus para
973 7 1.4 552 16 Q98C08 Q98C08 rhizobium l
974 7 1.4 553 5 Q9VVI0 Q9VVI0 drosophila
975 7 1.4 553 10 Q9FZL6 Q9FZL6 citrus sine
976 7 1.4 554 16 Q9RIS0 Q9RIS0 delnoccocus
977 7 1.4 555 16 Q9L022 Q9L022 streptomyce
978 7 1.4 557 16 Q9K4L0 Q9K4L0 streptomyce
979 7 1.4 559 17 Q9HPT5 Q9HPT5 halobacteri
980 7 1.4 560 16 Q9HVA3 Q9HVA3 pseudomonas
981 7 1.4 563 5 Q9VWJ2 Q9VWJ2 drosophila
982 7 1.4 564 16 Q9AA77 Q9AA77 caulobacter
983 7 1.4 565 3 Q13510 Q13510 rhodotorula
984 7 1.4 566 10 P77943 P77943 sulfolobus
985 7 1.4 566 10 Q93ZB4 Q93ZB4 arabisdopsis
986 7 1.4 566 16 P71838 P71838 mycobacteri
987 7 1.4 566 17 Q9C4U9 Q9C4U9 sulfolobus
988 7 1.4 568 5 Q9V600 Q9V600 drosophila
989 7 1.4 570 10 Q65813 Q65813 haematococc
990 7 1.4 570 16 P94544 P94544 bacillus su
991 7 1.4 571 5 Q95U47 Q95U47 drosophila
992 7 1.4 572 10 Q22056 Q22056 arabisdopsis
993 7 1.4 573 10 Q23212 Q23212 arabisdopsis
994 7 1.4 574 4 Q9H7K5 Q9H7K5 homo sapien
995 7 1.4 574 10 Q22455 Q22455 arabisdopsis
996 7 1.4 574 12 Q65563 Q65563 bovine herp
997 7 1.4 575 16 Q91642 Q91642 pseudomonas
998 7 1.4 578 5 Q9W324 Q9W324 drosophila
999 7 1.4 578 10 Q93Y74 Q93Y74 oryza sativ
1000 7 1.4 580 10 Q8VX08 Q8VX08 gentiana lu
```

## ALIGNMENTS

## RESULT 1

```
Q92797 PRELIMINARY: PRT: 493 AA.
ID Q92797
AC Q92797;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CT578 hypothetical protein.
GN CPN0809.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Ralman S., Mitchell W., Marathe R., Lammel C., Pan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RT Nat. Genet. 21:385-389(1999).
DR EMBL AE001662; AAD18947.1;
DR InterPro: IPR000508; SigPbase.
DR PROSITE: PS00501; SPASE_1.1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 493 AA: 50489 MW: 90F315B7A20AC2F0 CRC64;
```

Query Match 99.4%; Score 493; DB 16; Length 493;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 4 MSISSSGPDNKNINSQVLTSTPGVPOODKISGNETYQIOOTROGKNTMESDATTAG 63
DB 1 MSISSSGPDNKNINSQVLTSTPGVPOODKISGNETYQIOOTROGKNTMESDATTAG 60
```

```
QY 64 ASGKDKSTSTFKTERAPDOOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQTSI 123
DB 61 ASGKDKSTSTFKTERAPDOOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQTSI 120
QY 124 EASKSMESTLESLSLSAOKKEVEAVVVAALSGKSSGSAKLEPPELPKPGVTPRSEVI 183
DB 121 EASKSMESTLESLSLSAOKKEVEAVVVAALSGKSSGSAKLEPPELPKPGVTPRSEVI 180
QY 184 EIGLAKAKIOTLGEATKSNASTOQADQTNKLGLEKQAIKDKREPEYQEKAAE 243
DB 181 EIGLAKAKIOTLGEATKSNASTOQADQTNKLGLEKQAIKDKREPEYQEKAAE 240
QY 244 QSKDLEGTMDPVNTVMIVSAITVISTVAALFTGAGLAGLAAGAAAGAAAGAGA 303
DB 241 QSKDLEGTMDPVNTVMIVSAITVISTVAALFTGAGLAGLAAGAAAGAAAGAGA 300
QY 304 AATVVAQTITVQAVVQAVKQAVITAVROAITAIAKAASGKAFIKTLVNAIAKAISK 363
DB 301 AATVVAQTITVQAVVQAVKQAVITAVROAITAIAKAASGKAFIKTLVNAIAKAISK 360
QY 364 GISKVFAKGTOMIAKFKPLSKVYSSLTSKWTYVGVVVAAPALGKGIMOMLSMOON 423
DB 361 GISKVFAKGTOMIAKFKPLSKVYSSLTSKWTYVGVVVAAPALGKGIMOMLSMOON 420
QY 424 VAOFOKEVGKLAADMIAMFTQFMQASKIASKOTGESNEMQKATKLAGOILKAYAAI 483
DB 421 VAOFOKEVGKLAADMIAMFTQFMQASKIASKOTGESNEMQKATKLAGOILKAYAAI 480
QY 484 SGAIAGAKHTNNE 496
DB 481 SGAIAGAKHTNNE 493
```

## RESULT 2

```
Q9JS66 PRELIMINARY: PRT: 488 AA.
ID Q9JS66
AC Q9JS66;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CT578 hypothetical protein.
GN CPJ0809 OR Cpl062.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Nakata S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RT Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL: AE002263; AAF38835.1;
DR EMBL: AP002548; BAA99017.1;
DR TIGR: CP1062;
DR InterPro: IPR000508; SigPbase.
DR PROSITE: PS00501; SPASE_1.1; UNKNOWN_1.
SQ SEQUENCE 488 AA: 49819 MW: 06F2F0B905398AAB CRC64;
```









236	7	1.4	82	12	068295	068295 hepatitis c	309	7	1.4	108	2	08V1N7	08V1N7 pseudomonas
237	7	1.4	82	12	068296	068296 hepatitis c	310	7	1.4	108	3	09HRQ5	09HRQ5 candida alb
238	7	1.4	82	12	068297	068297 hepatitis c	311	7	1.4	108	5	09NN64	09NN64 leishmania
239	7	1.4	82	12	068298	068298 hepatitis c	312	7	1.4	109	2	09REM3	09REM3 streptomyc
240	7	1.4	82	12	068860	068860 hepatitis c	313	7	1.4	109	12	096861	096861 hepatitis c
241	7	1.4	82	12	068859	068859 hepatitis c	314	7	1.4	110	3	014317	014317 schizosacch
242	7	1.4	82	12	068947	068947 hepatitis c	315	7	1.4	110	9	08SDF3	08SDF3 pseudomonas
243	7	1.4	82	12	089948	089948 hepatitis c	316	7	1.4	110	12	09DMN7	09DMN7 hepatitis c
244	7	1.4	82	16	092RV1	092RV1 rhizobium m	317	7	1.4	110	12	09DMN6	09DMN6 hepatitis c
245	7	1.4	83	12	08V4A5	08V4A5 hepatitis c	318	7	1.4	110	12	09DMN0	09DMN0 hepatitis c
246	7	1.4	83	12	08V4A4	08V4A4 hepatitis c	319	7	1.4	110	12	09DMN9	09DMN9 hepatitis c
247	7	1.4	83	12	08V4A3	08V4A3 hepatitis c	320	7	1.4	110	12	09DMN6	09DMN6 hepatitis c
248	7	1.4	83	12	08V491	08V491 hepatitis c	321	7	1.4	110	12	09DMN3	09DMN3 hepatitis c
249	7	1.4	83	16	092LP2	092LP2 rhizobium m	322	7	1.4	110	12	09DML7	09DML7 hepatitis c
250	7	1.4	83	16	08V4C8	08V4C8 hepatitis c	323	7	1.4	110	12	09DML3	09DML3 hepatitis c
251	7	1.4	84	12	08V4C7	08V4C7 hepatitis c	324	7	1.4	110	12	09DMK8	09DMK8 hepatitis c
252	7	1.4	84	12	08V4C6	08V4C6 hepatitis c	325	7	1.4	110	12	09DMK5	09DMK5 hepatitis c
253	7	1.4	84	12	08V4C5	08V4C5 hepatitis c	326	7	1.4	110	12	09DMK6	09DMK6 hepatitis c
254	7	1.4	84	12	08V4C4	08V4C4 hepatitis c	328	7	1.4	111	3	09HGU9	09HGU9 aspergillus
255	7	1.4	84	12	08V4C3	08V4C3 hepatitis c	329	7	1.4	111	3	09HFO4	09HFO4 candida alb
256	7	1.4	84	12	08V4C1	08V4C1 hepatitis c	330	7	1.4	111	4	096QX6	096QX6 homo sapien
257	7	1.4	84	12	08V4C0	08V4C0 hepatitis c	331	7	1.4	111	11	099JDI1	099JDI1 rattus norv
258	7	1.4	84	12	08V4B9	08V4B9 hepatitis c	332	7	1.4	111	17	0976T2	0976T2 sulfolobus
259	7	1.4	84	12	08V4B8	08V4B8 hepatitis c	333	7	1.4	112	5	08WQJ0	08WQJ0 spodiopetra
260	7	1.4	84	12	08V4B7	08V4B7 hepatitis c	334	7	1.4	112	12	08I665	08I665 hepatitis c
261	7	1.4	84	12	08V4B6	08V4B6 hepatitis c	335	7	1.4	112	12	08I666	08I666 hepatitis c
262	7	1.4	84	12	08V4B5	08V4B5 hepatitis c	336	7	1.4	112	12	08I667	08I667 hepatitis c
263	7	1.4	84	12	08V4B4	08V4B4 hepatitis c	337	7	1.4	112	12	08I680	08I680 hepatitis c
264	7	1.4	84	12	08V4B3	08V4B3 hepatitis c	338	7	1.4	112	12	08I686	08I686 hepatitis c
265	7	1.4	84	12	08V4B2	08V4B2 hepatitis c	339	7	1.4	112	12	0127J9	0127J9 hepatitis c
266	7	1.4	84	12	08V4B1	08V4B1 hepatitis c	340	7	1.4	112	12	R87750	R87750 hepatitis c
267	7	1.4	84	12	08V4B0	08V4B0 hepatitis c	341	7	1.4	113	5	08SVZ9	08SVZ9 drosophila
268	7	1.4	84	12	08V4A9	08V4A9 hepatitis c	342	7	1.4	113	10	09ASX0	09ASX0 arabidopsis
269	7	1.4	84	12	08V4A8	08V4A8 hepatitis c	343	7	1.4	113	10	08S7W3	08S7W3 oryza sativ
270	7	1.4	84	12	08V4A7	08V4A7 hepatitis c	344	7	1.4	113	12	08I561	08I561 hepatitis c
271	7	1.4	84	12	08V4A6	08V4A6 hepatitis c	345	7	1.4	113	12	068835	068835 hepatitis c
272	7	1.4	84	12	08V4A5	08V4A5 hepatitis c	346	7	1.4	113	12	08I818	08I818 hepatitis c
273	7	1.4	84	12	08V4A1	08V4A1 hepatitis c	347	7	1.4	115	16	09L139	09L139 streptomyc
274	7	1.4	84	12	08V499	08V499 hepatitis c	348	7	1.4	118	16	0955D6	0955D6 drosophila
275	7	1.4	84	12	08V498	08V498 hepatitis c	349	7	1.4	118	16	09A7X0	09A7X0 caulobacter
276	7	1.4	84	12	08V497	08V497 hepatitis c	350	7	1.4	118	17	09Y9X0	09Y9X0 aeropyrum p
277	7	1.4	84	12	08V496	08V496 hepatitis c	351	7	1.4	122	2	P87108	P87108 escherichia
278	7	1.4	84	12	08V495	08V495 hepatitis c	352	7	1.4	125	2	08VTE2	08VTE2 thermus the
279	7	1.4	84	12	08V494	08V494 hepatitis c	353	7	1.4	126	2	08RP15	08RP15 pseudomonas
280	7	1.4	84	12	08V493	08V493 hepatitis c	354	7	1.4	126	2	09F1K8	09F1K8 streptomyc
281	7	1.4	84	12	08V492	08V492 hepatitis c	355	7	1.4	126	10	09XGP0	09XGP0 oryza sativ
282	7	1.4	84	12	08V490	08V490 hepatitis c	356	7	1.4	126	16	086327	086327 mycobacteri
283	7	1.4	84	12	08V460	08V460 hepatitis c	357	7	1.4	127	10	09LMD8	09LMD8 oryza sativ
284	7	1.4	84	12	08V459	08V459 hepatitis c	358	7	1.4	128	11	094DR8	094DR8 oryza sativ
285	7	1.4	84	12	08V458	08V458 hepatitis c	359	7	1.4	128	11	09CX69	09CX69 mus musculu
286	7	1.4	84	12	08V457	08V457 hepatitis c	360	7	1.4	128	11	08V7E8	08V7E8 tt virus. o
287	7	1.4	84	12	08V456	08V456 hepatitis c	361	7	1.4	128	12	08UYI0	08UYI0 tt virus. o
288	7	1.4	84	12	08V454	08V454 hepatitis c	362	7	1.4	130	17	09HNM8	09HNM8 halobacteri
289	7	1.4	84	12	08V453	08V453 hepatitis c	363	7	1.4	131	16	P73150	P73150 synecocyst
290	7	1.4	84	12	08V452	08V452 hepatitis c	364	7	1.4	133	1	074071	074071 cenarchaeum
291	7	1.4	84	12	08V451	08V451 hepatitis c	365	7	1.4	133	2	09RGK5	09RGK5 acinetobact
292	7	1.4	85	16	09A214	09A214 caulobacter	366	7	1.4	133	9	09MBP4	09MBP4 staphylococ
293	7	1.4	89	5	09NLS3	09NLS3 leishmania	367	7	1.4	133	12	091NQ4	091NQ4 tupaya hepp
294	7	1.4	90	16	09KYH3	09KYH3 streptomyc	368	7	1.4	134	10	09FRG4	09FRG4 oryza sativ
295	7	1.4	98	15	092Y48	092Y48 rhizobium m	369	7	1.4	134	12	09EB40	09EB40 hepatitis c
296	7	1.4	99	5	09B1I8	09B1I8 paragonimus	370	7	1.4	134	12	09EB39	09EB39 hepatitis c
297	7	1.4	100	11	09D475	09D475 mus musculu	371	7	1.4	134	12	09EB38	09EB38 hepatitis c
298	7	1.4	100	12	09WSC8	09WSC8 hepatitis c	372	7	1.4	134	12	09EB37	09EB37 hepatitis c
299	7	1.4	100	12	09WR65	09WR65 hepatitis e	373	7	1.4	134	12	09EB36	09EB36 hepatitis c
300	7	1.4	100	12	091TGO	091TGO hepatitis c	374	7	1.4	134	12	09EB35	09EB35 hepatitis c
301	7	1.4	100	16	08UC84	08UC84 agrobacteri	375	7	1.4	134	12	09EB27	09EB27 hepatitis c
302	7	1.4	101	10	0855Y9	0855Y9 oryza sativ	376	7	1.4	135	5	090875	090875 plasmodium
303	7	1.4	102	6	095KR1	095KR1 macaca fasc	377	7	1.4	137	5	09V5W8	09V5W8 drosophila
304	7	1.4	102	11	09CVY7	09CVY7 mus musculu	378	7	1.4	138	16	09A967	09A967 caulobacter
305	7	1.4	104	2	09RND9	09RND9 bordelella	379	7	1.4	140	12	068277	068277 hepatitis c
306	7	1.4	106	2	067984	067984 rhodococcus	380	7	1.4	140	12	068278	068278 hepatitis c
307	7	1.4	107	10	094DN9	094DN9 oryza sativ	381	7	1.4	140	12	068281	068281 hepatitis c
308	7	1.4	107	12	084703	084703 pea early b							



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:34:03 ; Search time 31 Seconds

(without alignments)  
3296.754 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496  
Sequence: 1 DTNMSISSSGPDNCKNINS.....LKAYAAISGALGAHKTNPF 496

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493	99.4	493	16	Q92797 chlamydia p
2	487	98.2	488	16	Q93666 chlamydia p
3	11	2.2	1937	5	Q9W313 drosophila
4	10	2.0	128	12	Q98279 molluscum c
5	10	2.0	137	15	Q36896 human immun
6	10	2.0	487	10	Q9P655 oryza sativ
7	10	1.8	103	10	Q8S076 oryza sativ
8	9	1.8	106	3	Q9HFC7 candida alb
9	9	1.8	109	3	Q8TFM9 fusarium cu
10	9	1.8	111	3	Q9C325 podospira a
11	9	1.8	128	12	Q8V7C9 lt virus. o
12	9	1.8	143	4	Q76068 homo sapien
13	9	1.8	147	10	Q94F07 arbidopsi
14	9	1.8	149	10	Q94F02 arbidopsi
15	9	1.8	153	10	Q9LY07 arbidopsi
16	9	1.8	153	10	Q94FR2 arbidopsi

Done  
Shane  
2/5/99

17	9	1.8	169	4	Q13850	Q13850 homo sapien
18	9	1.8	169	4	Q96TF6	Q96TF6 homo sapien
19	9	1.8	178	3	Q9UR55	Q9UR55 Kluyveromy
20	9	1.8	190	4	Q96LC3	Q96LC3 homo sapien
21	9	1.8	239	2	Q9S440	Q9S440 streptomyce
22	9	1.8	243	16	Q9S2Y8	Q9S2Y8 streptomyce
23	9	1.8	285	2	Q9RCT2	Q9RCT2 streptomyce
24	9	1.8	285	2	Q9RCT2	Q9RCT2 streptomyce
25	9	1.8	307	8	Q9XN22	Q9XN22 ascaris suu
26	9	1.8	308	4	Q9NPY2	Q9NPY2 homo sapien
27	9	1.8	474	16	Q8UJA1	Q8UJA1 agrobacteri
28	9	1.8	477	10	Q9SEF5	Q9SEF5 arbidopsi
29	9	1.8	487	16	Q84582	Q84582 chlamydia t
30	9	1.8	491	5	Q95TX7	Q95TX7 drosophila
31	9	1.8	491	16	Q9PJG2	Q9PJG2 chlamydia m
32	9	1.8	502	4	Q96LC4	Q96LC4 homo sapien
33	9	1.8	502	4	Q9BW38	Q9BW38 homo sapien
34	9	1.8	555	4	Q9NWM0	Q9NWM0 homo sapien
35	9	1.8	555	4	Q96CT3	Q96CT3 homo sapien
36	9	1.8	555	11	Q99K82	Q99K82 mus musculu
37	9	1.8	575	12	Q91TN6	Q91TN6 tupala herp
38	9	1.8	696	5	Q9VZY2	Q9VZY2 drosophila
39	9	1.8	746	10	Q9LJ00	Q9LJ00 arbidopsi
40	9	1.8	748	10	Q8W0N1	Q8W0N1 oryza sativ
41	9	1.8	810	11	Q9ES29	Q9ES29 mus musculu
42	9	1.8	950	2	Q50470	Q50470 mycobacteri
43	9	1.8	1013	5	Q96OM0	Q96OM0 drosophila
44	9	1.8	1017	16	Q96S86	Q96S86 mycobacteri
45	9	1.8	1017	5	Q96S86	Q96S86 mycobacteri
46	9	1.8	1024	5	Q9W268	Q9W268 drosophila
47	9	1.8	1475	5	Q9N4G4	Q9N4G4 caenorhabdi
48	9	1.8	1620	16	P96285	P96285 mycobacteri
49	9	1.8	2484	5	Q9U347	Q9U347 caenorhabdi
50	9	1.8	2607	5	Q92187	Q92187 caenorhabdi
51	9	1.8	2639	5	Q96786	Q96786 anthraea p
52	9	1.8	2655	5	Q964F4	Q964F4 anthraea y
53	9	1.6	26	4	Q9UPD8	Q9UPD8 gallus gall
54	9	1.6	60	13	Q92057	Q92057 gallus gall
55	8	1.6	67	3	Q94112	Q94112 pneumocysti
56	8	1.6	90	11	Q8R412	Q8R412 mus musculu
57	8	1.6	102	16	Q8U5Y9	Q8U5Y9 agrobacteri
58	8	1.6	106	3	Q94018	Q94018 candida alb
59	8	1.6	109	4	Q9N864	Q9N864 homo sapien
60	8	1.6	109	11	Q93J72	Q93J72 mus musculu
61	8	1.6	110	4	Q96U07	Q96U07 rhodotorula
62	8	1.6	129	4	Q43180	Q43180 homo sapien
63	8	1.6	140	10	Q9AY45	Q9AY45 oryza sativ
64	8	1.6	143	16	Q9HX45	Q9HX45 pseudomonas
65	8	1.6	145	16	Q9RRG3	Q9RRG3 delinococcus
66	8	1.6	157	10	Q9AMS9	Q9AMS9 oryza sativ
67	8	1.6	159	11	Q9C025	Q9C025 mus musculu
68	8	1.6	161	3	Q96WP4	Q96WP4 aspergillus
69	8	1.6	164	4	Q9NS18	Q9NS18 homo sapien
70	8	1.6	165	10	Q9AXF4	Q9AXF4 avicennia m
71	8	1.6	165	10	Q94G11	Q94G11 porteresia
72	8	1.6	165	10	Q94SE8	Q94SE8 pennicysti
73	8	1.6	166	3	Q9UVD0	Q9UVD0 pneumocysti
74	8	1.6	168	5	Q9NER1	Q9NER1 leishmania
75	8	1.6	178	5	Q45197	Q45197 caenorhabdi
76	8	1.6	178	16	Q9AAX0	Q9AAX0 caulobacter
77	8	1.6	179	10	Q9W0L8	Q9W0L8 arbidopsi
78	8	1.6	209	16	Q06630	Q06630 mycobacteri
79	8	1.6	212	10	Q9FTK1	Q9FTK1 oryza sativ
80	8	1.6	213	5	Q9USD5	Q9USD5 drosophila
81	8	1.6	213	5	Q18333	Q18333 drosophila
82	8	1.6	214	16	Q8UP60	Q8UP60 agrobacteri
83	8	1.6	223	16	Q9F2R7	Q9F2R7 streptomyce
84	8	1.6	225	2	Q8RNX8	Q8RNX8 propionibac
85	8	1.6	253	5	Q9BT14	Q9BT14 nephila sen
86	8	1.6	254	4	Q9H8P1	Q9H8P1 homo sapien
87	8	1.6	254	16	Q9XAT2	Q9XAT2 streptomyce
88	8	1.6	262	16	Q50464	Q50464 mycobacteri
89	8	1.6	273	2	Q9K188	Q9K188 bacillus ce



**THIS PAGE BLANK (USPTO)**

```

CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- DOMAIN: Highly repetitive protein characterized by regions of
CC polyalanine and glycine-rich repeating units.
CC -1- SIMILARITY: BELONGS TO THE SILK FIBROIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M37137; AAA29380.1; -
DR EMBL: U03848; AAB60212.1; -
DR PIR: A36068; A36068.
KW Silk Repeat.
FT NON_TER 1 1
FT DOMAIN 1 655
FT REPEAT 1 25
FT REPEAT 2 25
FT REPEAT 26 38
FT REPEAT 39 66
FT REPEAT 67 96
FT REPEAT 97 130
FT REPEAT 131 158
FT REPEAT 159 191
FT REPEAT 192 204
FT REPEAT 205 235
FT REPEAT 236 262
FT REPEAT 263 292
FT REPEAT 293 303
FT REPEAT 306 333
FT REPEAT 334 360
FT REPEAT 361 394
FT REPEAT 395 424
FT REPEAT 425 458
FT REPEAT 459 485
FT REPEAT 486 512
FT REPEAT 513 525
FT REPEAT 526 555
FT REPEAT 556 582
FT REPEAT 583 612
FT REPEAT 613 642
FT REPEAT 643 655
FT REPEAT 662 662
FT REPEAT 672 672
FT CONFLICT 695 747
FT CONFLICT 747 747
SQ SEQUENCE 747 AA; 60528 MW; 850E44BD649E012 CRC64;
      (IN REF. 1).
Query Match 1.6%; Score 8; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 293 GAAAGCA 300
      |||||
Db 624 GAAAGCA 631

```

```

RESULT 45
YNS4_CAEEL
ID YNS4_CAEEL STANDARD; PRT; 747 AA.
AC P34588;
DR 01-FEB-1994 (Rel. 28; Created)
DR 01-FEB-1994 (Rel. 28; Last sequence update)
DR 15-JUN-2002 (Rel. 41; Last annotation update)
DE Hypothetical protein ZC21.4 in chromosome III.
GN ZC21.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nemata; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol NZ;
RX MEDLINE=94150718; PubMed=7906398;
DR Wilson R., Ahncough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16685; AAA28171.1; ALT-INIT.
DR HSSP: Q07960; IRP.
DR WormPep: ZC21.4; CE00534.
DR InterPro: IPR00198; RhogAP.
DR Pfam: PF00620; RhogAP.
DR SMART: SM00324; RhogAP; 1.
KW SMART: hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 747 AA; 80829 MW; 48E9D61EDC4C56B0 CRC64;
      (IN REF. 1).
Query Match 1.6%; Score 8; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 299 GAAAGAAA 306
      |||||
Db 495 GAAAGAAA 502

```

Search completed: January 27, 2003, 16:35:16  
Job time : 35 secs

```

DR EMBL: AF054587; AAC08584.1; -.
DR Genew: HGNC:12968; ZNF179.
DR MIM: 601237; -.
DR InterPro: IPR001841; Znf_fing.
DR Pfam: PF00097; Zf-C3HC4; 1.
DR SMART: SM00184; RING_1.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
DR ZINC-finger.
KW ZNF-finger.
FT ZN-FING 57 98 RING-TYPE.
SQ SEQUENCE 632 AA; 68354 MW; F8D294815E65ED02 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 632;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAVGA 294
DB 584 AAGAAVGA 591

RESULT 43
CICL_RAT
ID CICL_RAT STANDARD; PRT; 687 AA.
AC P51802;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein CLC-K2.
GN CLCNKB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=9431614; PubMed=8041726;
RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;
RT "Two highly homologous members of the CLC chloride channel family in
RL both rat and human kidney.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=94292532; PubMed=8021279;
RA Adachi S., Uchida S., Hata M., Hirose M., Marumo F., Sasaki S.,
RA Ito H.;
RT "Two isoforms of a chloride channel predominantly expressed in thick
RL ascending limb of Henle's loop and collecting ducts of rat kidney.";
RL J. Biol. Chem. 269:17677-17683(1994).
CC -I- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY
CC CONCENTRATING MECHANISMS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CLC-K2L (SHOWN HERE) AND CLC-
CC K2S; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.
CC EXPRESSED IN ALL SEGMENTS OF THE NEPHRON EXAMINED, INCLUDING THE
CC S2 SEGMENT AND THE GLOMERULUS.
CC -I- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -I- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: Z30663; CAA83143.1; -.
DR EMBL: D26111; BAA05106.1; -.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001807; Cl-channel_volt.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00654; voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
DR SMART: SM00116; CBS; 1.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat; Alternative splicing.
FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 50 69 1 (POTENTIAL).
FT TRANSMEM 91 114 2 (POTENTIAL).
FT TRANSMEM 138 159 3 (POTENTIAL).
FT TRANSMEM 169 188 4 (POTENTIAL).
FT TRANSMEM 200 224 5 (POTENTIAL).
FT TRANSMEM 239 257 6 (POTENTIAL).
FT TRANSMEM 282 302 7 (POTENTIAL).
FT TRANSMEM 325 348 8 (POTENTIAL).
FT TRANSMEM 398 417 9 (POTENTIAL).
FT TRANSMEM 420 438 10 (POTENTIAL).
FT TRANSMEM 466 487 11 (POTENTIAL).
FT TRANSMEM 495 514 12 (POTENTIAL).
FT DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 646 664 13 (POTENTIAL).
FT DOMAIN 665 687 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 694 704 CBS 1.
FT DOMAIN 704 720 CBS 2.
FT CARBOHYD 364 364 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPIC 77 131 MISSING (IN ISOFORM CLC-K2S).
SQ SEQUENCE 687 AA; 75217 MW; 35B0BD775397880 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 LAAGAAVGA 293
DB 201 LAAGAAVGA 208

RESULT 44
SPDL_NEPCL
ID SPDL_NEPCL STANDARD; PRT; 747 AA.
AC P19837;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Spidroin 1 (dragline silk fibroin 1) (Fragment).
OS Nephila clavipes (Orb spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Melazoa; Arthropoda; Araneidae; Tetragnathidae; Nephila.
OX NCBI_TaxID=6915;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90384959; PubMed=2402494;
RA Xu M., Lewis R.V.;
RT "Structure of a protein superfiber: spider dragline silk.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:7120-7124(1990).
RN [2]
RP SEQUENCE OF 653-747 FROM N.A.
RX MEDLINE=94165058; PubMed=8120021;
RA Beckwith R., Arcidiacono S.;
RT "Sequence conservation in the C-terminal region of spider silk
RL proteins (Spidroin) from Nephila clavipes (Tetragnathidae) and
RL Araneus bicentenarius (Araneidae).";
RL J. Biol. Chem. 269:6661-6663(1994).
CC -I- FUNCTION: Spiders major ampullate silk possesses unique
CC characteristics of strength and elasticity. Fibroin consists of
CC pseudocrystalline regions of antiparallel beta-sheet interspersed
CC with elastic amorphous segments.
CC -I- SUBUNIT: MAJOR SUBUNIT, WITH SPIDROIN 2, OF THE DRAGLINE SILK.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC3.
RX MEDLINE=94064614; PubMed=8244983;
RA Gott J.M., Visomirski L.M., Hunter J.L.;
RT "Substitutional and insertional RNA editing of the cytochrome c
RL oxidase subunit 1 mRNA of Physarum polycephalum.";
J. Biol. Chem. 268:25483-25486(1993).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme. Electrons originating in
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme a of subunit 1 to the binuclear center formed by heme a3
and copper B.
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) -> 4 ferricytochrome
c + 2 H2O.
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane. Contains 12 potential transmembrane domains.
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
CC -1- CAUTION: SOME POSITIONS ARE MODIFIED BY RNA EDITING.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14769; AAA71993.1; -.
DR HSSP; P18401; 1FFT.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
FM Respiratory chain; Inner membrane; RNA editing.
FT METAL 87 87 IRON (HEME A) (PROBABLE).
FT METAL 266 266 COPPER B (PROBABLE).
FT METAL 270 270 COPPER B (PROBABLE).
FT METAL 315 315 COPPER B (PROBABLE).
FT METAL 316 316 COPPER B (PROBABLE).
FT METAL 401 401 IRON (HEME A3) (PROBABLE).
FT METAL 403 403 IRON (HEME A) (PROBABLE).
SQ SEQUENCE 594 AA; 66513 MW; 8D8A60CA4620D5FF CRC64;

Query Match 1.6%; Score 8; DB 1; Length 594;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLA 287
DB 49 GAGLAGLA 56

RESULT 41
2179_RAT STANDARD: PRT; 631 AA.
AC 2179_RAT
AC 070418;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein 179 (Brain finger protein).
GN ZNF179 OR BFP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=98042506; PubMed=9367872;

```

```

RA Inoue S., Oriimo A., Saito T., Ikeda K., Sakata K., Hosoi T., Oriimo H.,
RA Ouchi Y., Muramatsu M.;
RT "A novel RING finger protein, BFP, predominantly expressed in the
RT brain.";
RL Biochem. Biophys. Res. Commun. 240:8-14(1997).
CC -1- TISSUE SPECIFICITY: Predominantly expressed in brain.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF054586; AAC08583.1; -.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Zinc-finger.
FT ZN_FING 57 98 RING-TYPE.
SQ SEQUENCE 631 AA; 68646 MW; 6C997B8890CAB564 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAVGA 294
DB 583 AAGAAVGA 590

RESULT 42
2179_HUMAN STANDARD: PRT; 632 AA.
AC 09ULX5; O60633;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein 179 (Brain finger protein).
GN ZNF179 OR BFP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Saito T.;
RT "Human BFP/ZNF179.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
[2]
DE SEQUENCE OF 33-127 FROM N.A.
RX MEDLINE=96301415; PubMed=8660987;
RA Matsuda Y., Inoue S., Seki N., Hosoi T., Oriimo A., Muramatsu M.,
RA Hori T.;
RT "Chromosome mapping of human (ZNF179), mouse, and rat genes for brain
RL finger protein (bfp), a member of the RING finger family.";
CC Genomics 33:325-327(1996).
CC -1- TISSUE SPECIFICITY: Predominantly expressed in brain.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB026054; BAA84698.1; -.
DR

```

Query Match 1.6%; Score 8; DB 1; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 354 VKAIKAI 361  
 Db 57 VKAIKAI 64

```

RESULT 38
LEU1_LISMO STANDARD; PRT; 512 AA.
ID LEU1_LISMO
AC O8Y5R9;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 2-isopropylmalate synthase (EC 4.1.3.12) (Alpha-isopropylmalate
  synthase) (Alpha-IPM synthetase).
GN LEUA OR LMO1987.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
  Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
  Charbit A., Cheouani F., Couve E., de Daruvar A., Depoux P.,
  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
  Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
  Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
  Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
  Notides G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
  Remmel B., Rose M., Schlueder T., Simoes N., Tierrez A.,
  Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
  *Comparative genomics of Listeria species.*;
  Science 294:849-852(2001).
RL
RT
CC -1- FUNCTION: Catalyzes the condensation of the acetyl group of
  acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
  3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
CC -1- CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + CoA =
  acetyl-CoA + 3-methyl-2-oxobutanoate + H(2O).
CC -1- PATHWAY: Leucine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE
  SYNTHASE FAMILY. LEUA 1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591981; CAD00065.1; -.
CC Listlist: LMO01987; -.
DR InterPro: IPR002034; AIPM/Hct-synth.
DR InterPro: IPR000891; HMG-L-like.
DR Pfam: PF00682; HMG-L-like; 1.
DR TIGRfams: TIGR00973; leua_bact; 1.
DR PROSITE: PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
DR PROSITE: PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 512 AA; 56096 MW; 5AABID633DAB791 CRC64;

```

Query Match 1.6%; Score 8; DB 1; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 354 VKAIKAI 361

Db 57 VKAIKAI 64

```

RESULT 39
Y4TO_RHISN STANDARD; PRT; 531 AA.
ID Y4TO_RHISN
AC P55669;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Probable peptide ABC transporter periplasmic binding protein Y4TO
  precursor.
GN Y4TO.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
  Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
  Perret X.;
  *Molecular basis of symbiosis between Rhizobium and legumes.*;
  Nature 387:394-401(1997).
RL
RT
CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
  SYSTEM Y4TOFORs FOR A PEPTIDE.
CC -1- SUBCELLULAR LOCATION: Periplasmic (potential).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
  PROTEIN FAMILY 5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000098; AAB91868.1; -.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
DR PROSITE: PS01040; SBP_BACTERIAL_5; 1.
KW Hypothetical protein; Transport; Periplasmic; Signal; Plasmid.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 531 PROBABLE PEPTIDE ABC TRANSPORTER
  PERIPLASMIC BINDING PROTEIN Y4TO.
SQ SEQUENCE 531 AA; 58138 MW; F2C1EB3EBB48B27 CRC64;

```

Query Match 1.6%; Score 8; DB 1; Length 531;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 284 AGLAGAA 291  
 Db 11 AGLAGAA 18

```

RESULT 40
COX1_PHYPO STANDARD; PRT; 594 AA.
ID COX1_PHYPO
AC C07434;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Cytochrome c oxidase polypeptide I (bc 1.9.3.1).
GN COX1.
OS Physarum polycephalum (Slime mold).
OC Mitochondrion.
OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
  Physarum.
OX NCBI_TaxID=5791;

```

```

DR PIR: S20939; S20939.
DR InterPro: IPR002937; Amino_oxidase.
DR Pfam: PF01593; Amino_oxidase; 1.
KW Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.
KM Membrane; Herbicide resistance; Complete proteome.
FT NP_BIND 7 23 FAD (ADP PART) (POTENTIAL).
FT VARIANT 195 195 R -> C (CONFERES RESISTANCE TO THE
SEQUENCE 472 AA; 52920 MW; 81D0896DAA28758 CRC64;
HERBICIDE NORFLURAZON).
OY 280 GAGIAGLA 287
Db 7 GAGIAGLA 14
|||||
RESULT 36
ZDS_ANASP STANDARD; PRT; 479 AA.
ID ZDS_ANASP
AC Q9R6X4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zeta-carotene desaturase (EC 1.14.99.30) (Carotene 7,8-desaturase).
GN CRRQ OR ALL2382.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RP SEQUENCE FROM N.A.
RA Mann V., Hirschberg J.;
RT "Evolution of the carotenoid biosynthesis pathway: two pathways for
lycopene synthesis exist in Anabaena PCC7120."
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
RA Matnabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RT DNA Res. 8:205-213(2001).
CC -1- FUNCTION: Catalyzes the conversion of zeta-carotene to lycopene
via the intermediary of neurosporene. It carries out two
consecutive desaturations (introduction of double bonds) at
positions C-7 and C-7' (By similarity).
CC -1- CATALYTIC ACTIVITY: Zeta-carotene + AH(2) + O(2) = neurosporene +
A + 2 H(2)O.
CC -1- CATALYTIC ACTIVITY: Neurosporene + AH(2) + O(2) = lycopene + A + 2
H(2)O.
CC -1- COFACTOR: NAD, NADP, OR FAD (PROBABLE).
CC -1- PATHWAY: Carotenoid biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ZETA CAROTENE DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y15115; CAB56041.1; -.
DR EMBL: AP003589; BAB74081.1; -.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR002050; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.

```

```

KW Carotenoid biosynthesis; Oxidoreductase; NAD; Flavoprotein; FAD;
KM Complete proteome.
FT CONFILCT 92 MISSING (IN REF. 1).
SO SEQUENCE 479 AA; 53579 MW; D8B14BD098965F01 CRC64;
OY 280 GAGIAGLA 287
Db 7 GAGIAGLA 14
|||||
RESULT 37
LEU1_LISTIN STANDARD; PRT; 512 AA.
ID LEU1_LISTIN
AC Q92A28;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-Isopropylmalate synthase (EC 4.1.3.12) (Alpha-isopropylmalate
synthase) (Alpha-IPM synthetase).
GN LEUOA OR LIN2094.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnuk C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chablit A., Chetoui P., Couve E., de Darvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Ertlan K.-P., Esbel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel M., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tietze A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:845-852(2001).
CC -1- FUNCTION: Catalyzes the condensation of the acetyl group of
acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
CC -1- CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + CoA =
acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O.
CC -1- PATHWAY: Leucine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE
SYNTHASE FAMILY. LEUOA 1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596171; CAC97324.1; -.
DR EMBL: AL596171; CAC97324.1; -.
DR InterPro: IPR002034; AIPM/HcIt_synth.
DR InterPro: IPR000891; HMGL-like.
DR Pfam: PF00682; HMGL-like; 1.
DR TIGRfams: TIGR00973; leuA_bact; 1.
DR PROSITE: PS00815; AIPM_HOMOCITR_SYNTH_1; 1.
DR PROSITE: PS00816; AIPM_HOMOCITR_SYNTH_2; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
SO SEQUENCE 512 AA; 56149 MW; 23836636815C782 CRC64;

```

DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Molybdenum cofactor synthase protein 3 (Molybdopterin synthase  
 DE sulfurylase) (MPT synthase sulfurylase).  
 GN MOC53.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Scallmeyer B., Coyne K.E., Wuebbers M.M., Johnson J.L.,  
 RA Rajagopalan K.V., Mendel R.R.;  
 RT "The cDNA sequence of MOC53, human molybdopterin synthase  
 RT sulfurylase.";  
 RN Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.R., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lovell J.D.,  
 RA Leharvalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swan R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=uterus;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Activates MPT synthase by the ATP dependant adenylation  
 CC of its C-terminal residue (By similarity).  
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
 CC HESG/MOE/THIF FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF102544; AAC72412.1; -  
 CC EMBL: AB034553; CAB53750.1; -  
 CC EMBL: BC015939; AAH15939.1; -  
 CC GeneW: HGNC:15765; MOC53.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR InterPro: IPR001763; Rhodanese-like.

DR InterPro: IPR000594; Thif\_domain.  
 DR Pfam: PF00581; Rhodanese; 1.  
 DR Pfam: PF00899; Thif; 1.  
 DR SMART: SM00450; RHOD; 1.  
 DR KEGG: Molybdenum cofactor biosynthesis.  
 FT DOMAIN 350 460 RHODANESE.  
 SQ SEQUENCE 460 AA; 49669 MW; 29944E75513EE324 CRC64;  
 Query Match 1.68; Score 8; DB 1; Length 460;  
 Best local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 137 LQSLAAO 144  
 1111111111  
 425 LQSLAAO 432  
 RESULT 35  
 CRT1-SYNY3  
 ID CRT1-SYNY3 STANDARD: PRT: 472 AA.  
 AC P29273;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).  
 DE PDS OR CRD OR SDR1254.  
 GN Synchocystis sp. (strain PCC 6803).  
 OS Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
 OC NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92256820; PubMed=1581575;  
 RA Martinez-Ferez I.M., Vioque A.;  
 RT "Nucleotide sequence of the phytoene desaturase gene from  
 RT Synchocystis sp. PCC 6803 and characterization of a new mutation  
 RT which confers resistance to the herbicide norflurazon.";  
 RL Plant Mol. Biol. 18:981-983(1992).  
 RN [2]  
 RP REVISIONS TO C-TERMINUS.  
 RA Vioque A.;  
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
 RA Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE  
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO  
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.  
 CC -1- COFACTOR: NAD, NADP, OR FAD (PROBABLE).  
 CC -1- ENZYME REGULATION: INHIBITED BY THE HERBICIDE NORFLURAZON IN A  
 CC NON-COMPETITIVE WAY.  
 CC -1- PATHWAY: Carotenoid biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X62574; CAA44452.1; -  
 CC EMBL: D90909; BAA17847.1; -

```

FT TRANSMEM 14 34 POTENTIAL.
SQ SEQUENCE 344 AA; 36421 MW; AC37A4EF919B71B5 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAGAA 301
DB 49 AAGAA 56

RESULT 32
HYPR_ALCEU STANDARD; PRT; 361 AA.
AC P31902;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hydrogenase nickel incorporation protein hypr.
GN HYPR.
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H16 / DSM 428 / ATCC 17699;
RX MEDLINE=93356597; PubMed=8352644;
RA Derrenede J., Ellinger M., Friedrich B.;
RT "Analysis of a pleiotropic gene region involved in formation of
catalytically active hydrogenases in Alkaligenes eutrophus H16.";
RT Arch. Microbiol. 153:345-353(1993).
RN [2]
RP REVISIONS.
RC STRAIN-H16 / DSM 428 / ATCC 17699;
RA Ellinger J.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COULD BE INVOLVED IN NICKEL BINDING AND ACCUMULATION.
CC -1- SIMILARITY: BELONGS TO THE HYPR/HUPM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70183; CAA49732.1; -
DR PIR; S29976;
DR InterPro; IPR004392; Hypr.
DR InterPro; IPR002894; Hypr_Ureg.
DR Pfam; PF01495; Hypr_Ureg_1
DR TIGRfams; TIGR00073; Hypr_2;
KW Metal-binding; Nickel; Plasmid.
SQ SEQUENCE 361 AA; 38483 MW; BC42976B7C5F7C4 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 300 AAGAAAT 307
DB 323 AAGAAAT 330

RESULT 33
TRMU_MYCTU STANDARD; PRT; 367 AA.
ID TRMU_MYCTU
AC 053271;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE (EC 2.1.1.61).
DE TRMU OR RV3024C OR MT3108 OR MTV012.39C.
GN Mycobacterium tuberculosis.
OS Mycobacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Bacteria; Actinobacteria; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=96295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RT Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RT Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
homocysteine + tRNA containing 5-methylaminomethyl-2-
thiouridylate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TRMU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL021287; CAM16109.1; -
DR EMBL; AE007129; AK47438.1; -
DR TIGR; MT3108; -
DR Tuberculist; RV3024C; -
DR InterPro; IPR004506; Trmu.
DR InterPro; IPR004135; tRNA_Me_trans.
DR Pfam; PF03054; tRNA_Me_trans_1.
DR TIGRfams; TIGR00420; trmu; 1.
KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 367 AA; 38120 MW; 043ECC2D5D690C79 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 ATTAGAG 66
DB 351 ATTAGAG 358

RESULT 34
MOC3_HUMAN STANDARD; PRT; 460 AA.
ID MOC3_HUMAN
AC 095396;

```



```

DE Lipopolysaccharide core biosynthesis protein rfas.
GN Rfas OR B3629.
OS Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92325066; PubMed=1624461;
RA Pridel E., Parker C.T., Schmittman C.A.;
RT "Structures of the rfaB, rfaI, rfaJ, and rfaS genes of Escherichia
RT coli K-12 and their roles in assembly of the lipopolysaccharide
RT core."
RL J. Bacteriol. 174:4736-4745(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes."
RL Nucleic Acids Res. 22:2576-2586(1994).
CC -i- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80599; AAA24084.1; -.
DR EMBL: U00039; AB18606.1; -.
DR EMBL: AE000440; AAC7653.1; -.
DR PIR: A42982; A42982.
DR ECGene: EG1350; rfas.
KW Lipopolysaccharide biosynthesis; Complete proteome.
SQ SEQUENCE 311 AA; 36730 MW; C03F1BBE6BDC26 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 KDKTSSTT 74
DB 191 KDKTSSTT 198

RESULT 30
PSBO_EUGR STANDARD: PRT: 338 AA.
ID PSBO_EUGR
AC P46483;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 1, chloroplast precursor (OEEL).
GN PSBO.
OS Euglena gracilis.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 94-110.
RC STRAIN=Z;
RX MEDLINE=94154233; PubMed=8111018;
RA Shigemori Y., Inagaki J., Mori H., Nishimura M., Takahashi S.,
RA Yamamoto Y.;
RT "The presence of the precursor to the nucleus-encoded 30 kDa
RT protein of photosystem II in Euglena gracilis Z includes two
RT hydrophobic domains."
RL Plant Mol. Biol. 24:209-215(1994).
CC -i- FUNCTION: STABILIZES THE MANGANESE CLUSTER WHICH IS THE PRIMARY

```

```

CC SITE OF WATER SPLITTING (BY SIMILARITY).
CC -i- SUBUNIT: MONOMER.
CC -i- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -i- SIMILARITY: BELONGS TO THE PSBO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14702; BAA03529.2; ALT_INT.
DR InterPro: IPR002628; PSI1_MSP.
DR Pfam: PF01716; MSP. 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transist peptide;
KW Thylakoid; Membrane; Manganese.
FT TRANSIT 1 93 CHLOROPLAST.
FT CHAIN 94 338 OXYGEN-EVOLVING ENHANCER PROTEIN 1.
SQ SEQUENCE 338 AA; 35271 MW; 1034FE4010B1A2B CRC64;

Query Match 1.6%; Score 8; DB 1; Length 338;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 AGAANGAA 295
DB 41 AGAANGAA 48

RESULT 31
LIC2_BURCE STANDARD: PRT: 344 AA.
ID LIC2_BURCE
AC Q92EM5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipase chaperone (Lipase foldase) (Lipase helper protein)
DE (Lipase activator protein) (Lipase modulator).
GN LIPB OR HP.
OS Burkholderia cepacia (pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 21808;
RX MEDLINE=99124623; PubMed=9925617;
RA Ouyen D.T., Schmidt-Dannert C., Schmid R.D.;
RT "High-level formation of active pseudomonas cepacia lipase after
RT heterologous expression of the encoding gene and its modified
RT chaperone in Escherichia coli and rapid in vitro refolding."
RL Appl. Environ. Microbiol. 65:787-794(1999).
CC -i- FUNCTION: SEEMS TO BE ACTING IN THE FOLDING OF THE EXTRACELLULAR
CC LIPASE DURING ITS PASSAGE THROUGH THE PERIPLASM.
CC -i- SUBCELLULAR LOCATION: INNER-MEMBRANE ANCHORED (BY SIMILARITY).
CC -i- SIMILARITY: BELONGS TO THE LIPASE CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ131766; CAA10510.1; -.
DR InterPro: IPR004961; Lipase_chap.
DR Pfam: PF03280; Lipase_chap. 1.
KW Lipid degradation; Chaperone; Transmembrane; Periplasmic;
KW Inner membrane.

```

```

CC (VCFS/DCS). BELONGS TO THE PAIRED HOMEOBOX FAMILY.
CC -1- SIMILARITY: BELONGS TO THE PAIRED HOMEOBOX FAMILY.
CC "BICOID" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U96402; AAC39544.1; -.
DR HSSP: P06601; 1FJL.
DR TRANSFAC: T04038; -.
DR Genew: HGNC:4613; GSCl.
DR MIM: 601845; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS50071; HOMEOBOX_2; 1.
DR Homeobox; DNA-binding; Nuclear protein; Polymorphism.
FT DOMAIN 2 POLY-ALA.
FT DNAS_BIND 64 70 POLY-CYS.
FT VARIANT 126 185 HOMEOBOX.
FT VARIANT 47 47 R->C.
SQ SEQUENCE 205 AA; 21545 MW; 665C33D9C454E7A7 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 294 AAAAGGAA 301
DB 2 AAAAGGAA 9
RESULT 27
RL10_EUGGR STANDARD; PRT; 215 AA.
AC Q39724;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10.
GN RPL10.
OS Euglena gracilis.
OC Eukaryota; Euklenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Deruere J., Schantz M.L., Schantz R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L10F FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94087; CAA63831.1; -.
DR InterPro: IPR001197; Ribosomal_L10E.
DR Pfam: PF00826; Ribosomal_L10e; 1.
DR TIGRFAMs: TIGR00279; L10e; 1.
DR PROSITE: PS01257; RIBOSOMAL_L10e; 1.
DR Ribosomal protein.

```

```

SQ SEQUENCE 215 AA; 24642 MW; 4B329C0C13F37B2E CRC64;
Query Match 1.6%; Score 8; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 149 EAVVVAAL 156
DB 145 EAVVVAAL 152
RESULT 28
VG67_HSVB STANDARD; PRT; 272 AA.
AC P28984;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Hypothetical gene 67 protein (IR6 protein).
GN 67 OR IR6.
OS Equine herpesvirus type 1 (strain AB4P) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520; 10329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB4P.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RL "The DNA sequence of equine herpesvirus-1."
RT Virology 189:304-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Kentucky A;
RX MEDLINE=93079867; PubMed=1333117;
RA Breiden C.A., Yalamanchili R.R., Colle C.F., III, O'Callaghan D.J.;
RT "Identification and transcriptional mapping of genes encoded at the
RL IR/US junction of equine herpesvirus type 1."
RN Virology 191:649-660(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86664; AAB02512.1; -.
DR EMBL: M86664; AAB02502.1; -.
DR EMBL: M80429; AAA46076.1; -.
DR PIR: D36802;
DR PIR: B44215; B44215.
KW Hypothetical protein.
SQ SEQUENCE 272 AA; 30103 MW; 448000315F376654 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 294 AAAAGGAA 301
DB 39 AAAAGGAA 46
RESULT 29
RFAS_ECOLI STANDARD; PRT; 311 AA.
AC P27126;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```
CC -----
CC EMBL: M7332; AAA82712.1; .
CC PIR: A40814; A40814.
CC InterPro: IPR002379; ATPase_Csub.
CC InterPro: IPR00245; Vac_ATPsyn_Csub.
CC Pfam: PF00137; ATP-synt_C; 2.
CC PRINTS: PR00122; VACATPASE.
CC TIGRFAMS: TIGR01100; V_ATP_synth_C; 1.
CC DR TIGRFAMS: TIGR01100; V_ATP_synth_C; 1.
CC KW Hydrolyase: Hydrogen ion transport; ATP synthesis; Transmembrane;
CC Multigene family.
CC FT DOMAIN 1 10 LUMENAL (POTENTIAL).
CC FT TRANSMEM 11 33 POTENTIAL.
CC FT DOMAIN 34 55 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 56 76 POTENTIAL.
CC FT DOMAIN 77 95 LUMENAL (POTENTIAL).
CC FT TRANSMEM 96 117 POTENTIAL.
CC FT DOMAIN 118 129 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 130 155 POTENTIAL.
CC FT DOMAIN 156 165 LUMENAL (POTENTIAL).
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC FT VARIANT 78 78 T -> P (IN CLONE 12).
CC FT VARIANT 89 89 A -> V (IN CLONE 93).
CC FT VARIANT 164 164 A -> V (IN CLONE 93).
CC SQ SEQUENCE 165 AA; 16621 MW; 900A664C6C1965BB CRC64;

Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 165;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 GIAGLAG 289
DB 103 GIAGLAG 110

RESULT 25
VATL_ORYSA STANDARD; PRT; 165 AA.
AC 040635;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).
GN VAMP-P1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
RN NM_117111;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Guang-1u-al No.4; TISSUE=Shoot;
RA Xiao C.;
RL Thesis (1995), Fudan University, China.
CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
CC EUKARYOTIC CELLS.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN, WHICH IS PRESENT
CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: U27098; AAA68175.1; .
CC InterPro: IPR002379; ATPase_Csub.
CC InterPro: IPR00245; Vac_ATPsyn_Csub.
CC Pfam: PF00137; ATP-synt_C; 2.
CC PRINTS: PR00122; VACATPASE.
CC DR TIGRFAMS: TIGR01100; V_ATP_synth_C; 1.
CC KW Hydrolyase: Hydrogen ion transport; ATP synthesis; Transmembrane.
CC FT DOMAIN 1 10 LUMENAL (POTENTIAL).
CC FT TRANSMEM 11 33 POTENTIAL.
CC FT DOMAIN 34 55 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 56 76 POTENTIAL.
CC FT DOMAIN 77 95 LUMENAL (POTENTIAL).
CC FT TRANSMEM 96 117 POTENTIAL.
CC FT DOMAIN 118 129 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 130 155 POTENTIAL.
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC SQ SEQUENCE 165 AA; 16667 MW; E580168BDD0FD1 CRC64;

Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 165;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 GIAGLAG 289
DB 103 GIAGLAG 110

RESULT 26
GSC_L_HUMAN STANDARD; PRT; 205 AA.
ID GSC_L_HUMAN
AC 015499;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein goosecoid-like (GSC-2).
GN GSC_L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN NM_117111;
RP SEQUENCE FROM N.A., AND VARIANT CYS-47.
RX MEDLINE=98110571; PubMed=9441739;
RA Funke B., St Jore B., Puech A., Strokin H., Edelman L., Carlson C.,
RA Raft S., Pandita R.K., Kucherlapati R., Skoultschi A., Morrow B.E.;
RT "Characterization and mutation analysis of goosecoid-like (GSC_L), a
RT homeodomain-containing gene that maps to the critical region for
RT VCFS/DCS on 22q11."
RT Genomics 46:364-372(1997).
RN NM_117111;
RP SEQUENCE FROM N.A.
RX MEDLINE=97294411; PubMed=9150167;
RA Gottlieb S., Emanuel B.S., Driscoll D.A., Sellinger B., Wang Z.,
RA Roe B., Budarf M.L.;
RT "The digeorge syndrome minimal critical region contains a goosecoid-
RT like (GSC_L) homeobox gene that is expressed early in human
RT development."
RT Am. J. Hum. Genet. 60:1194-1201(1997).
CC -1- FUNCTION: MAY HAVE A ROLE IN DEVELOPMENT. MAY REGULATE ITS OWN
CC TRANSCRIPTION. MAY BIND THE BTCD01 CONSENSUS SEQUENCE TATATC.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: DETECTED IN ADULT TESTIS AND PITUITARY, AND IN
CC 9-10 WK FETAL TISSUE (THORAX). PROBABLY EXPRESSED IN OTHER TISSUES
CC AT LOW LEVELS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EARLY HUMAN DEVELOPMENT AS WELL
CC AS IN A LIMITED NUMBER OF ADULT TISSUES.
CC -1- DISEASE: MAY BE RELATED TO SOME OF THE DEVELOPMENTAL DEFECTS
CC ASSOCIATED WITH VELOCARDIOFACIAL AND DIGEORGE SYNDROMES
```

AC P50515; (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).  
 GN VMA3 OR SPAC183.14.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetes; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 ON NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92206078; PubMed=1839480;  
 RA Toyama R., Goldstein D.J., Schlegel R., Dhar R.;  
 RT "A genomic sequence of the Schizosaccharomyces pombe 16 kDa vacuolar  
 RT H(+)-ATPase."  
 RL Yeast 7:989-991(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=972;  
 RA MEDLINE=2184401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H.,  
 RA Bozzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaude V., Motlier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,  
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potshkin J.,  
 RA Shpukovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE  
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE  
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN  
 CC EUKARYOTIC CELLS.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN: WHICH IS PRESENT  
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.  
 CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYLOHEXYLCARBODIIMIDE (DCCD)  
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X5947; CAA42572.1; -

DR EMBL; Z98598; CAB1240.1; -  
 DR InterPro: IPR002379; ATPase\_Csub.  
 DR InterPro: IPR000245; Vac\_ATPsyn\_Csub.  
 DR Pfam: PF00137; ATP-synt\_C; 2.  
 DR PRINTS: PR00122; VACATPASE.  
 DR TIGRFS: TIGR01100; V\_ATP\_synt\_C; 1.  
 KW Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.  
 FT DOMAIN 1  
 FT TRANSMEM 10 32 POTENTIAL.  
 FT DOMAIN 33 54 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 55 75 POTENTIAL.  
 FT DOMAIN 76 91 LUMENAL (POTENTIAL).  
 FT TRANSMEM 92 113 POTENTIAL.  
 FT DOMAIN 114 125 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 126 151 POTENTIAL.  
 FT DOMAIN 152 161 LUMENAL (POTENTIAL).  
 FT BINDING 138 138 DICYLOHEXYLCARBODIIMIDE (POTENTIAL).  
 SQ SEQUENCE 161 AA; 16325 MW; DEF52DFDC8CEB39 CRC64;  
 Query Match 1.6%; Score 8; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 7.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 282 GLAGLAG 289  
 DB 99 GLAGLAG 106  
 ID VATL\_AVEA STANDARD; PRT; 165 AA.  
 AC P23957;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).  
 GN VATP-P1.  
 OS Avena sativa (Oat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Avenae; Avena  
 ON NCBI\_TaxID=4498;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANG; PubMed=1831453;  
 RX MEDLINE=91340758;  
 RA Lai S., Watson J.C., Hansen J.N., Sze H.;  
 RT "Molecular cloning and sequencing of cDNAs encoding the proteolipid  
 RT subunit of the vacuolar H(+)-ATPase from a higher plant."  
 RL J. Biol. Chem. 266:16078-16084(1991).  
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE  
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE  
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN  
 CC EUKARYOTIC CELLS.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN: WHICH IS PRESENT  
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.  
 CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYLOHEXYLCARBODIIMIDE (DCCD)  
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X5947; CAA42572.1; -

KW Hypothetical protein; Membrane; Lipoprotein; Signal;  
 KW Complete proteome.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 132 HYPOTHETICAL LIPOPROTEIN AQ.615.  
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).  
 FT SEQUENCE 132 AA; 13815 MM; A303B3203A67F0AD CRC64;  
 SO  
 Query Match 1.6%; Score 8; DB 1; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 289 GAAVGAAA 296  
 DB 35 GAAVGAAA 42  
 RESULT 21  
 NRDI\_CORAM  
 ID NRDI\_CORAM STANDARD; PRT; 144 AA.  
 AC 069272;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nrdi protein.  
 GN NRDI.  
 OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).  
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 CC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;  
 CC Corynebacterium.  
 CC NCBI\_TaxID=1697;  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 6872;  
 RX MEDLINE=96136125; PubMed=9468481;  
 RA Fieschi F., Torrents E., Toulkhouva L., Jordan A., Hellman U.,  
 RA Barde J., Gilbert I., Karlsson M., Sjoberg B.M.;  
 RA "The manganese-containing ribonucleotide reductase of Corynebacterium  
 RA ammoniagenes is a class Ib enzyme.";  
 RL J. Biol. Chem. 273:4329-4337(1998).  
 CC -1- FUNCTION: NOT KNOWN. PROBABLY INVOLVED IN RIBONUCLEOTIDE REDUCTASE  
 CC FUNCTION.  
 CC -1- SIMILARITY: BELONGS TO THE NRDI FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y09572; CAA70764.1; -;  
 DR InterPro; IPR004465; NRDI.  
 DR TrEMBL; TIGR00333; nrdi; 1.  
 SO SEQUENCE 144 AA; 15897 MM; FEID345DB629D251 CRC64;  
 Query Match 1.6%; Score 8; DB 1; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 6.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 219 KIGLEKQA 226  
 DB 136 KIGLEKQA 143  
 RESULT 22  
 VATL\_NEUCR  
 ID VATL\_NEUCR STANDARD; PRT; 161 AA.  
 AC P31413;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).

GN VMA-3 OR 12F11.130.  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariales; Sordariaceae; Neurospora.  
 CC NCBI\_TaxID=5141;  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=93015814; PubMed=1400281;  
 RX Bowman B.J., Sista H.;  
 RA "The vacuolar ATPase of Neurospora crassa.";  
 RL J. Bioenerg. Biomembr. 24:361-370(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A;  
 RA Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.-W., Mannhaupt G.;  
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE  
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE  
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN  
 CC EUKARYOTIC CELLS.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN, WHICH IS PRESENT  
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.  
 CC -1- SUBCELLULOUS: THIS SUBUNIT BINDS DICYCLOHEXYLCARBODIIMIDE (DCDD)  
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L07105; AAA19974.1; -;  
 DR EMBL; AL451017; CAC18222.1; -;  
 DR InterPro; IPR002379; ATPase\_Csub.  
 DR InterPro; IPR000245; Vac\_ATPsyn\_Csub.  
 DR Pfam; PF00137; ATP\_synth\_C; 2.  
 DR PRINTS; PRO0122; VACATPASE.  
 DR TrEMBL; TIGR01100; V\_ATP\_synth\_C; 1.  
 KW Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.  
 FT DOMAIN 1 8 LUMENAL (POTENTIAL).  
 FT TRANSMEM 9 31 LUMENAL (POTENTIAL).  
 FT DOMAIN 32 53 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 54 74 LUMENAL (POTENTIAL).  
 FT DOMAIN 75 91 LUMENAL (POTENTIAL).  
 FT TRANSMEM 92 113 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 114 125 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 126 151 LUMENAL (POTENTIAL).  
 FT DOMAIN 152 161 LUMENAL (POTENTIAL).  
 FT BINDING 138 138 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).  
 SO SEQUENCE 161 AA; 16329 MM; D7C54CD209373EB4 CRC64;  
 Query Match 1.6%; Score 8; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 7.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 282 GIAGLAAG 289  
 DB 99 GIAGLAAG 106  
 RESULT 23  
 VATL\_SCHPO  
 ID VATL\_SCHPO STANDARD; PRT; 161 AA.

OY 294 AAAAGGA 301  
 DB 72 AAAAGGA 79

## RESULT 18

RLA3\_CLAHE STANDARD: PRT: 111 AA.

ID RLA3\_CLAHE  
 AC P42038;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 60S acidic ribosomal protein p2 (Allergen Cla h 3) (Cla h III).  
 GN CLAH3.  
 OS Cladosporium herbarum.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;  
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;  
 OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.  
 OX NCBI\_TaxID=29918;

SEQUENCE FROM N.A.  
 RA MEDLINE=95114396; PubMed=7814877;  
 RX Zhang L., Muradia G., Curran I.H., Rode H., Vijay H.M.;  
 RT "A cDNA clone coding for a novel allergen, Cla h III, of Cladosporium  
 herbarum identified as a ribosomal p2 protein.";  
 RT J. Immunol. 154:710-717(1995).  
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF  
 CC PROTEIN SYNTHESIS.  
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL  
 CC SUBUNIT.  
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -1- CAUTION: Two distinct proteins have been termed allergen  
 CC Cla h 3.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X77253; CAA54470.1; -  
 CC InterPro: IPR001813; 60S\_ribosomal; 1.  
 CC Pfam: PF00428; 60S\_ribosomal; 1.  
 CC Ribosomal protein; Phosphorylation; Multigene family; Allergen.  
 KW SEQUENCE 111 AA; 11124 MW; 297FA13FEBEDF89 CRC64;  
 SQ

Query Match 1.6%; Score 8; DB 1; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 5.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGA 300  
 DB 78 GAAAGGA 85

## RESULT 19

RLA3\_TRYCR STANDARD: PRT: 112 AA.

ID RLA3\_TRYCR  
 AC P26795;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE 60S acidic ribosomal protein p2-B (P2B).  
 OS Trypanosoma cruzi.  
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RA;

MEDLINE=92310999; PubMed=1614880;

RA Vazquez M.P., Schlijman A.G., Panhebra A., Levin M.J.;  
 RT "Nucleotide sequence of a cDNA encoding another Trypanosoma cruzi  
 RT acidic ribosomal p2 type protein (TcP2B)."  
 RL Nucleic Acids Res. 20:2893-2893(1992).  
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF  
 CC PROTEIN SYNTHESIS.  
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL  
 CC SUBUNIT.  
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X65065; CAA46198.1; -  
 DR PIR: S22950; R6UT2B.  
 DR InterPro: IPR001813; 60S\_ribosomal; 1.  
 DR Pfam: PF00428; 60S\_ribosomal; 1.  
 KW Ribosomal protein; Phosphorylation.  
 SQ SEQUENCE 112 AA; 10926 MW; DE84C739B4F048BF CRC64;

Query Match 1.6%; Score 8; DB 1; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 5.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGA 301  
 DB 81 AAAAGGA 88

## RESULT 20

Y615\_AQUAE STANDARD: PRT: 132 AA.

ID Y615\_AQUAE  
 AC O66867;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical lipoprotein AQ\_615 precursor.  
 GN AQ\_615.

OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
 OC Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.

STRAIN=VF5;  
 MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M., Keller M., Anjaj M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus.";  
 RL Nature 392:353-358(1998).

CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL: AE000698; AAC06830.1; -  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 CC

ID RLA4\_SCHPO STANDARD; PRT: 110 AA.  
AC P17478;  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 60S acidic ribosomal protein p2-beta (A4).  
GN RPP2B OR RPA4 OR SPBC23G7.15C.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
XX [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90220620; PubMed=2325655;  
RA Beltrame M., Bianchi M.E.;  
RT "A gene family for acidic ribosomal proteins in Schizosaccharomyces  
RL pombe: two essential and two nonessential genes.";  
RM Mol. Cell. Biol. 10:2341-2348(1990).  
RN [2]  
PP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Squeros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voicikert G., Aert R., Robben J., Grymptre B.,  
RA Welljans I., Vastreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Medler H., Mambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,  
RA Domínguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF  
CC PROTEIN SYNTHESIS.  
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL  
CC SUBUNIT.  
CC -1- PM: PHOSPHORYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A  
CC PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR  
CC BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO  
CC ONE OF THE TWO A PROTEINS PRESENT IN MULTICELLULAR ORGANISMS.  
CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL.  
CC WHEREAS RPA1 AND RPA2 ARE NOT.  
CC -1- SIMILARITY: ALL FOUR RPA SHOW 35% IDENTITY. RPA4 IS HIGHLY RELATED  
CC TO RPA2.  
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M33142; AAA53337.1; -

DR EMBL; AL035065; CAA22631.1; -  
DR PIR; D34715; R6BY24.  
DR InterPro: IPR001813; 60S\_ribosomal.  
DR Pfam; PF00428; 60S\_ribosomal; 1.  
DR Rfam; RF00428; 60S\_ribosomal; 1.  
KW Ribosomal protein; Phosphorylation; Multigene family.  
SQ SEQUENCE 110 AA; 11120 MW; 477D7B04CF36520CF CRC64;  
  
Query Match 1.6%; Score 8; DB 1; Length 110;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 296 AAGGAGAGA 303  
Db 75 AAGGAGAGA 82  
  
RESULT 17  
RLA2\_ARTSA STANDARD; PRT: 111 AA.  
ID RLA2\_ARTSA  
AC P02399;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE 60S acidic ribosomal protein p2 (E112).  
OS Artemia salina (Brine shrimp).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
OC Branchiopoda; Anostraca; Artemiidae; Artemia.  
OX NCBI\_TaxID=85549;  
XX [1]  
RP SEQUENCE OF 1-108 FROM N.A.  
RX MEDLINE=85230659; PubMed=3839187;  
RA Maassen J.A., Schop E.N., Brands J.H.G.M., van Hemert F.J.,  
RA Lenstra J.A., Møller W.;  
RT "Molecular cloning and analysis of cDNA sequences for two ribosomal  
RT proteins from Artemia. The coordinate expression of genes for  
RT ribosomal proteins and elongation factor 1 during embryogenesis of  
RT Artemia.";  
RL Eur. J. Biochem. 149:609-616(1985).  
RN [2]  
PP SEQUENCE.  
RX MEDLINE=80004136; PubMed=477981;  
RA Amos R., Pluijms W.J.M., Moeller W.;  
RT "The primary structure of ribosomal protein e12/e12-p from Artemia  
RT salina 80 S ribosomes.";  
RL FEBS Lett. 104:85-89(1979).  
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF  
CC PROTEIN SYNTHESIS.  
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL  
CC SUBUNIT.  
CC -1- PM: PHOSPHORYLATED (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X02632; CAA26479.1; -  
DR PIR; A25208; R8S812.  
DR InterPro: IPR001813; 60S\_ribosomal.  
DR Pfam; PF00428; 60S\_ribosomal; 1.  
KW Ribosomal protein; Phosphorylation.  
FT MOD\_RES 98 98 PHOSPHORYLATION (PARTIAL).  
FT CONFLICT 19 19 S -> T (IN REF. 2).  
FT CONFLICT 80 80 T -> A (IN REF. 2).  
SQ SEQUENCE 111 AA; 11503 MW; 69C4241E1FAF4E01 CRC64;  
  
Query Match 1.6%; Score 8; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X92375; CAA63119.1; -
DR EMBL: X92374; CAA63118.1; -
DR EMBL: M95063; AAA18550.1; -
DR InterPro: IPR002379; ATPase_Csdb.
DR Pfam: PF00137; ATP-synt_C; 1.
DR TRIGREMS: TRIGR01100; V_ATP_synt_C; 1.
KM Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.
FT NON_TER 1 1
FT TRANSMEM <1 20 POTENTIAL.
FT TRANSMEM 21 39 LUMENAL (POTENTIAL).
FT TRANSMEM 40 61 POTENTIAL.
FT TRANSMEM 62 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 99 POTENTIAL.
FT DOMAIN 100 109 LUMENAL (POTENTIAL).
FT BINDING 86 86 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
FT CONFLICT 52 52 A -> P (IN REF. 1).
FT CONFLICT 63 63 A -> G (IN REF. 1); CAA63119).
SQ SEQUENCE 109 AA; 11043 MW; 6B7518DD7A0FEB62 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GIAGIAG 289
Db 47 GIAGIAG 54

RESULT 15
RLA2_SCHPO STANDARD; PRT: 110 AA.
AC P08094;
AC 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein P2-alpha (A2) (L40C) (L12E1).
GN RPP2A OR RPA2 OR SPB8B7.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88067727; PubMed=3684587;
RX Beltrame M., Bianchi M.E.;
RT "Sequence of the cDNA for one acidic ribosomal protein of
RT Schizosaccharomycetes pombe."
RL Nucleic Acids Res. 15:9089-9089(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220620; PubMed=2325655;
RX Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
RT pombe: two essential and two nonessential genes."
RL Mol. Cell. Biol. 10:2341-2348(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RX Squires J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RX Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

```

```

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver R., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds S., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gilmoprre B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Pohl T.M.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Laure V., Motter S.,
RA Gallibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Talada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen O., Potashkin J.,
RA Snopcewski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomycetes pombe."
RA Nature 415:871-880(2002).
[4]
RP SEQUENCE OF 1-40.
RX MEDLINE=84038947; PubMed=6355773;
RX Otake E., Higo K.-H., Itoh T.;
RT "Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from
RT Schizosaccharomycetes pombe."
RL Mol. Gen. Genet. 191:519-524(1983).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A
CC PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR
CC BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO
CC ONE OF THE TWO A PROTEINS PRESENT IN MULTICELLULAR ORGANISMS.
CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL,
CC WHEREAS RPA1 AND RPA2 ARE NOT.
CC -1- SIMILARITY: ALL FOUR RPA SHOW 35% IDENTITY. RPA4 IS HIGHLY RELATED
CC TO RPA2.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00466; CAA68528.1; -
DR EMBL: M3138; AAA35335.1; -
DR EMBL: AL032684; CAA21791.1; -
DR PIR: B34715; R6BY22.
DR InterPro: IPR001813; 60S_ribosomal.
DR Pfam: PF00428; 60S_ribosomal.1.
KW Ribosomal protein; Phosphorylation; Multigene family.
SQ SEQUENCE 110 AA; 11158 MW; 2B465A46E5160FE CRC64;

Query Match 1.6%; Score 8; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGAA 301
Db 76 AAAAGAA 83

RESULT 16
RLA4_SCHPO

```



RT "ORK1, a potassium-selective leak channel with two pore domains  
RT cloned from *Drosophila melanogaster* by expression in *Saccharomyces*  
RT *cerevisiae*.";  
RT Proc. Natl. Acad. Sci. U.S.A. 93:13256-13261(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abiri J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dev I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunlov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Malteli B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissensbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
RL Science 287:2185-2195(2000).  
CC -1- FUNCTION: BACKGROUND POTASSIUM CHANNEL. RECTIFICATION IS DEPENDENT  
CC ON EXTERNAL POTASSIUM CONCENTRATION. ACTS AS AN OUTWARDLY  
CC RECTIFYING CHANNEL BUT AS EXTERNAL POTASSIUM LEVELS INCREASE, THIS  
CC IS REVERSED.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN ADULT, STRONGEST  
CC EXPRESSION IN MUSCLE, BRAIN AND OVARY. ALSO PRESENT AT LOW LEVELS  
CC IN LARVA AND EMBRYO.  
CC -1- MISCELLANEOUS: INHIBITED BY BARIUM.  
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM  
CC CHANNELS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U55321; AAC69250.1; -  
CC EMBL: AE003484; AAF47972.1; -  
CC FLYBase: FBgn0017561; ORK1.  
CC InterPro: IPR003280; K+channel\_2pore.  
CC InterPro: IPR001622; K+channel\_pore.  
CC InterPro: IPR000636; M+channel\_nlg.  
CC Pfam: PF00520; Ion\_trans; 1.

DR PRINTS: PRO1333; 2PORECHANNEL.  
KW Ionic channel; Transmembrane; Ion transport; Potassium transport;  
KW Glycoprotein.  
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 7 27 POTENTIAL.  
FT DOMAIN 95 111 PORE-FORMING 1 (POTENTIAL).  
FT TRANSSEM 120 140 POTENTIAL.  
FT DOMAIN 141 170 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 171 191 POTENTIAL.  
FT DOMAIN 208 224 PORE-FORMING 2 (POTENTIAL).  
FT TRANSSEM 244 264 POTENTIAL.  
FT DOMAIN 265 1001 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 58 58 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 1001 AA; 109289 MW; 09AE1A369072E07 CRC64;  
Query Match 1.88; Score 9; DB 1; Length 1001;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 294 AAAGGAG 302  
Db 764 AAAGGAG 772  
RESULT 14  
VATL\_MAIZE STANDARD: PRT: 109 AA.  
ID VATL\_MAIZE  
AC Q41773; Q41774; Q08074;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14) (V-  
DE ATPase 16 kDa proteolipid subunit) (Fragment).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE OF 1-76 FROM N.A.  
RC STRAIN=cv. Lixis; TISSUE=coleoptile;  
RX MEDLINE=96197807; PubMed=8617373;  
RA Viereck R., Kirsch M., Loew R., Rausch T.;  
RT "Down-regulation of plant V-type H<sup>+</sup>-ATPase genes after light-induced  
RT inhibition of growth.";  
RT FEBS Lett. 384:285-288(1996).  
RN [2]  
RP SEQUENCE OF 52-109 FROM N.A.  
RC STRAIN=cv. B73; TISSUE=leaf;  
RX MEDLINE=94105294; PubMed=8278499;  
RA Keith C.S., Hoang D.O., Barrett B.M., Felgelman B., Nelson M.C.,  
RA "Partial sequence analysis of 130 randomly selected maize cDNA  
RT clones.";  
RL Plant Physiol. 101:329-332(1993).  
CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE  
CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE  
CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN  
CC EUKARYOTIC CELLS.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +  
CC H(+)(Out).  
CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN, WHICH IS PRESENT  
CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.  
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE MESOCOTYL TIP OF  
CC ETIOLATED SEEDLINGS COMPARED TO THE BASE.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS STRONGLY LINKED TO EXTENSION  
CC GROWTH.  
CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYCLONEXYLICARBOXYLIMIDE (DCDI)  
CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).

KW Capsid assembly.  
 SO SEQUENCE 782 AA; 87392 MW; 4DE40314E5BAB9B CRC64;  
 Query Match 1.8%; Score 9; DB 1; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGAAA 301  
 |||||||||  
 DB 436 GAAAGAAA 444

RESULT 11  
 ELS\_MOUSE STANDARD; PRT; 860 AA.  
 AC P54320;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Elastin precursor (tropoelastin).  
 GN ELN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE=Lung;  
 RX MEDLINE=95130069; PubMed=7829060;  
 RA Wyder K.S., Sechler J.L., Boyd C.D., Passmore H.C.:  
 RT "Use of an intron polymorphism to localize the tropoelastin gene to  
 mouse chromosome 5 in a region of linkage conservation with human  
 chromosome 7".  
 RT Genomics 23:125-131(1994).  
 CC -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND  
 CC NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.  
 CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER  
 CC INTO AN EXTENSIBLE 3D NETWORK.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.  
 CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; 008210; AAA80155.1; -  
 DR MGSD; MG1:95317; Eln.  
 DR InterPro: IPR003979; tropoelastin.  
 DR PRINTS: PR01500; TROPOELASTIN.  
 KW Structural protein; Repeat; Signal; Connective tissue.  
 FT SIGNAL 1 27  
 FT CHAIN 28 860 ELASTIN.  
 FT SEQUENCE 860 AA; 71955 MW; 0C0BE5AAE1EDD7F1 CRC64;  
 SO

Query Match 1.8%; Score 9; DB 1; Length 860;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 GAAAGAAA 306  
 |||||||||  
 DB 113 GAAAGAAA 121

RESULT 12  
 SOXA\_COR81 STANDARD; PRT; 967 AA.  
 ID SOXA\_COR81  
 AC 046337;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sarcosine oxidase alpha subunit (EC 1.5.3.1) (Sarcosine oxidase  
 subunit).  
 GN SOXA.  
 OS Corynebacterium sp. (strain P-1).  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;  
 OC Corynebacterium.  
 OX NCBI\_TaxID=69006;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-P-1;  
 RX MEDLINE=95355441; PubMed=7543100;  
 RA Chlumsky L.J., Zhang L., Jorns M.S.:  
 RT "Sequence analysis of sarcosine oxidase and nearby genes reveals  
 RT homologues with key enzymes of folate one-carbon metabolism.";  
 RL J. Biol. Chem. 270:18252-18259(1995)  
 CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEMETHYLATION OF SARCOSE TO  
 CC YIELD GLYCINE, HYDROGEN PEROXIDE AND 5,10-  
 CC METHYLENETHETRAHYDROFOLATE.  
 CC -1- CATALYTIC ACTIVITY: Sarcosine + H(2)O + O(2) = glycine +  
 CC formaldehyde + H(2)O(2).  
 CC -1- COFACTOR: FAD.  
 CC -1- SUBUNIT: HETEROTETRAMER (ALPHA 100 kDa, BETA 42 kDa, GAMMA 20 kDa,  
 CC AND DELTA 6 kDa).  
 CC -1- SIMILARITY: BELONGS TO THE GCVT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; 023955; AAC43461.1; -  
 DR InterPro: IPR002536; GCV\_T.  
 DR InterPro: IPR000205; NAD-binding.  
 DR Pfam: PF01571; GCV\_T; 1.  
 DR ProDom: PD000139; FAD-pyr\_redox; 1.  
 KW Oxidoreductase; Flavoprotein; FAD.  
 SO SEQUENCE 967 AA; 102767 MW; C5F8A61F021134E9 CRC64;  
 QY 289 GAAVAAAA 297  
 |||||||||  
 DB 434 GAAVAAAA 442

RESULT 13  
 ORK1\_DROME STANDARD; PRT; 1001 AA.  
 ID ORK1\_DROME  
 AC 094526;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Open rectifier potassium channel protein 1 (Two pore domain potassium  
 channel ORK1).  
 GN ORK1 OR CG1615.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Larva;  
 RX MEDLINE=97075152; PubMed=8917578;  
 RA Goldstein S.A.N., Price L.A., Rosenthal D.N., Pausch M.H.;

Db 15 GAGIAGLAA 23

RESULT 9

ELS\_BOVIN STANDARD; PRT; 747 AA.

ID ELS\_BOVIN STANDARD; PRT; 747 AA.

AC P04985; P04986; P04987; Q29421;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Elastin precursor (Tropoelastin).

GN ELN.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.

OC NCBI\_TaxID=9913;

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87194772; PubMed=3032943;

RA Raju K., Anwar R.A.;

RT "Primary structures of bovine elastin a, b, and c deduced from the sequences of cDNA clones."

RL J. Biol. Chem. 262:5755-5762(1987).

RN [2]

RP SEQUENCE OF 1-27 FROM N.A.

RX TISSUE-Nuchal ligament:

RA MEDLINE=89274159; PubMed=2543440;

RA Yeh H., Anderson N., Ornstein-Goldstein N., Bashir M.M., Mecham R., Rosenbloom J.C., Abrams W.R., Indik Z., Yoon K., Parks W.,

RA Rosenbloom J.;

RT "Structure of the bovine elastin gene and S1 nuclease analysis of alternative splicing of elastin mRNA in the bovine nuchal ligament."

RL Biochemistry 28:2365-2370(1989).

RN [3]

RP SEQUENCE OF 1-27 FROM N.A.

RX MEDLINE=91234332; PubMed=2031719;

RA Manohar A., Shi W., Anwar R.A.;

RT "Partial characterization of bovine elastin gene; comparison with the gene for human elastin."

RL Biochem. Cell Biol. 69:185-192(1991).

CC -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.

CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER INTO AN EXTENSIBLE 3D NETWORK.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: J02717; AAA30503.1; -

DR EMBL: K03505; AAA30505.1; -

DR EMBL: K03506; AAA30506.1; -

DR EMBL: J02855; AAA30776.1; -

DR EMBL: M58652; AAA03519.2; -

DR PIR: A26728; A26728.

DR PIR: B26728; B26728.

DR PIR: C26728; C26728.

DR InterPro: IPR003979; tropoelastin.

DR PRINTS: PR01500; TROPOELASTIN.

KW Structural protein; Connective tissue; Repeat; Signal;

KW Alternative splicing

FT SIGNAL 1 26

FT CHAIN 27 747 ELASTIN.

FT MOD\_RES 105 105 OXIDATIVE DEAMINATION.

FT MOD\_RES 109 109 OXIDATIVE DEAMINATION.

FT MOD\_RES 252 252 OXIDATIVE DEAMINATION.

FT MOD\_RES 271 271 OXIDATIVE DEAMINATION.

FT MOD\_RES 275 275 OXIDATIVE DEAMINATION.

FT MOD\_RES 324 324 OXIDATIVE DEAMINATION.

FT MOD\_RES 327 327 OXIDATIVE DEAMINATION.

FT MOD\_RES 400 400 OXIDATIVE DEAMINATION.

FT MOD\_RES 404 404 OXIDATIVE DEAMINATION.

FT MOD\_RES 407 407 OXIDATIVE DEAMINATION.

FT MOD\_RES 448 448 OXIDATIVE DEAMINATION.

FT MOD\_RES 489 489 OXIDATIVE DEAMINATION.

FT MOD\_RES 493 493 OXIDATIVE DEAMINATION.

FT MOD\_RES 544 544 OXIDATIVE DEAMINATION.

FT MOD\_RES 548 548 OXIDATIVE DEAMINATION.

FT MOD\_RES 552 552 OXIDATIVE DEAMINATION.

FT MOD\_RES 606 606 OXIDATIVE DEAMINATION.

FT MOD\_RES 609 609 OXIDATIVE DEAMINATION.

FT MOD\_RES 645 645 OXIDATIVE DEAMINATION.

FT MOD\_RES 649 649 OXIDATIVE DEAMINATION.

FT MOD\_RES 685 685 OXIDATIVE DEAMINATION.

FT MOD\_RES 688 688 OXIDATIVE DEAMINATION.

FT MOD\_RES 226 239 MISSING (IN ELASTIN B).

FT VARSPLIC 226 259 MISSING (IN ELASTIN C).

FT CONFLICT 1 3 MRS -> MAG (IN REF. 2 AND 3).

FT CONFLICT 12 12 E -> G (IN REF. 2 AND 3).

FT SEQUENCE 747 AA; 64229 MW; 633C03E411643D83 CRC64;

Query Match 1.8%; Score 9; DB 1; Length 747;

Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 298 GGAGAGAAA 306

Db 95 GGAGAGAAA 103

RESULT 10

ID PRTP\_HSVT2 STANDARD; PRT; 782 AA.

AC O9WR16;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable processing and transport protein.

GN PRTP.

OS Herpesvirus tupaia (Strain 2) (THV-2).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Betaherpesvirinae.

OC NCBI\_TaxID=132678;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99319892; PubMed=10392721;

RA Bahn U., Springfield C., Tidona C.A., Darai G.;

RT "Structural organization of a conserved gene cluster of Tupaia herpesvirus encoding the DNA polymerase, glycoprotein B, a probable processing and transport protein, and the major DNA binding protein."

RL Virus Res. 60:123-136(1999).

CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLLOCATION OF THE VIRUS GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION (BY SIMILARITY).

CC -----

CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AF084543; AAD42934.1; -

DR InterPro: IPR000501; Proc\_transport.

DR Pfam: PF01366; PRTP; 1.

TRANSMEMBRANE SEGMENTS M1, M2, AND M4 FROM SUBUNITS OF THE GABA(A)/BENZODIAZEPINE RECEPTOR FAMILY, RESPECTIVELY.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: M36035; AAA03652.1; -  
 EMBL: U12421; AAA83252.1; -  
 EMBL: Z83214; CAB55884.1; -  
 EMBL: BC001110; AAH01110.1; -  
 PIR: S14257; S14257.  
 Genew: HGNC:1138; B2RP.  
 MIM: 109610; -  
 InterPro: IPR004307; Tspo\_MBR.  
 Pfam: PF03073; Tspo\_MBR; 1.  
 Mitochondrion: Receptor; Transmembrane; Polymorphism.

DR TRANSMEM 5 26 TM1 (POTENTIAL).  
 KM TRANSMEM 47 67 TM2 (POTENTIAL).  
 FT TRANSMEM 80 100 TM3 (POTENTIAL).  
 FT TRANSMEM 106 126 TM4 (POTENTIAL).  
 FT TRANSMEM 135 155 TM5 (POTENTIAL).  
 FT TRANSMEM 147 147 A -> T.  
 FT VARIANT 162 162 /FTid=VAR\_013617.  
 FT VARIANT 162 162 H -> R.  
 FT VARIANT 162 162 /FTid=VAR\_013618.  
 FT VARIANT 162 162 IAD741BF9AB92CD CRC64;  
 FT SEQUENCE 169 AA; 18779 MW; 1AD741BF9AB92CD CRC64;

Query Match 1.8%; Score 9; DB 1; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310  
 DB 117 GAAATTVA 125

RESULT 7  
 CRTI\_STRGR STANDARD; PRT; 507 AA.  
 AC P54981; P72447;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).  
 GN CRTI OR CRTI.  
 OS Streptomyces griseus.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1911;  
 RN NCB1\_TaxID=1911;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JA3933;  
 RX MEDLINE=97074881; PubMed=8917308;  
 RA Schumann G., Nurnberger H., Sandmann G., Kruegel H.J.;  
 RT "Activation and analysis of cryptic crt genes for carotenoid biosynthesis from Streptomyces griseus.";  
 RT Mol. Gen. Genet. 252:656-666(1996).  
 RL -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.  
 CC -1- COFACTOR: FAD (PROBABLE).  
 CC -1- PATHWAY: Carotenoid biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: L37405; AAA1950.1; -  
 EMBL: X95396; CAA64850.1; -  
 DR EMBL: X95396; CAA64850.1; -  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR00171; Bac\_phytoene\_dh.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PROSITE: PS00982; PHYTOENE\_DH; 1.  
 KW Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.  
 FT NP\_BIND 12 45 FAD (ADP PART) (POTENTIAL).  
 FT NP\_BIND 507 AA; 54509 MW; FBB97FEE696B2AC CRC64;  
 FT SEQUENCE 507 AA; 54509 MW; FBB97FEE696B2AC CRC64;

Query Match 1.8%; Score 9; DB 1; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288  
 DB 15 GAGLAGLAA 23

RESULT 8  
 CRTI\_STRGR STANDARD; PRT; 508 AA.  
 AC P54971;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).  
 GN CRTI.  
 OS Streptomyces setonii.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=38315;  
 RN NCB1\_TaxID=38315;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ISP 5395;  
 RA Hoshii K.;  
 RT Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.  
 RL -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.  
 CC -1- COFACTOR: FAD (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: D55723; BAA09537.1; -  
 DR EMBL: D55723; BAA09537.1; -  
 DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR00171; Bac\_phytoene\_dh.  
 DR InterPro: IPR00205; NAD\_binding.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR PROSITE: PS00982; PHYTOENE\_DH; 1.  
 KW Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.  
 FT NP\_BIND 12 45 FAD (ADP PART) (POTENTIAL).  
 FT NP\_BIND 508 AA; 54610 MW; 4B6DEFC076D51CB5 CRC64;  
 FT SEQUENCE 508 AA; 54610 MW; 4B6DEFC076D51CB5 CRC64;

Query Match 1.8%; Score 9; DB 1; Length 508;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288  
 DB 15 GAGLAGLAA 23

```

RESULT 5
RLA4 CLAHE          STANDARD:          PRT:      111 AA.
ID P42039;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein p2 (Minor allergen Cla h 4) (Cla h IV).
GN CLA4..
OS Cladosporium herbarum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
OX NCBI_TaxID=29918;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=280202-Berlin;
RX MEDLINE=95206305; PubMed=7898496;
RA Achatz G., Oberkofler H., Lechenauer E., Simon B., Unger A.,
RA Kandler D., Ehner C., Prillinger H., Kraft D., Breitenbach M.;
RT "Molecular cloning of major and minor allergens of Alternaria
RT alternata and Cladosporium herbarum.";
RL Mol. Immunol. 32:213-227(1995).
RN [2]
RP REVISIONS TO 13; 38-41 AND 93.
RA Simon-Nobbe B.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DDAJ databases.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- CAUTION: Two distinct proteins have been termed allergen
CC Cla h 4.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
-----
DR EMBL: X78223; CAA55067.2; -
DR InterPro: IPR001813; 60S_ribosomal.
DR Pfam: PF00428; 60S_ribosomal; 1.
KW Ribosomal protein; Phosphorylation; Multigene family; Allergen.
SQ SEQUENCE 111 AA; 11105 MW; C7B65CGAD97876A CRC64;

Query Match      1.8%; Score 9; DB 1; Length 111;
Best Local Similarity 100.0%; Pred No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 293 GAAAGGAA 301
   |||||
Db 77 GAAAGGAA 85

RESULT 6
PKBS_HUMAN          STANDARD:          PRT:      169 AA.
ID P30516; Q96TF6;
AC 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peripheral-type benzodiazepine receptor (PBR) (PKBS) (Mitochondrial
DE benzodiazepine receptor).
GN BZRP OR MBR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146565; PubMed=1847678;
RA Rioud J., Mattei M.-G., Kagnad M., Dumont X., Guillemot J.C.,
RA le Fur G., Caput D., Ferrara P.;
RT "Molecular cloning and chromosomal localization of a human
RT peripheral-type benzodiazepine receptor.";
RL Eur. J. Biochem. 195:305-311(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95237610; PubMed=7721091;
RA Yakovlev A.G., Ruffo M., Jurka J., Krueger K.E.;
RT "Comparison of repetitive elements in the third intron of human and
RT rodent mitochondrial benzodiazepine receptor-encoding genes.";
RL Gene 155:201-205(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M., Lloyd C., Lloyd D.M.,
RA Marlyn I.D., Mashregh-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillips B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Steward C.A., Sultson J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaali S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White T., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chissee S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Kort I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Salita S., Budarf M.L.,
RA Modermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim Y.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiliann Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDAJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR THE MANIFESTATION OF PERIPHERAL-TYPE
CC BENZODIAZEPINE RECOGNITION SITES AND IS MOST LIKELY TO COMPRISE
CC BINDING DOMAINS FOR BENZODIAZEPINES AND ISOQUINOLINE CARBOAMIDES.
CC MAY PLAY A ROLE IN THE TRANSPORT OF PORPHYRINS AND HEME.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL; INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: SEGMENTS TMI, TM4, AND TM5 SHOW SIMILARITY WITH THE

```

```

FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 344 364 POTENTIAL.
FT TRANSMEM 377 397 POTENTIAL.
FT TRANSMEM 409 429 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
FT TRANSMEM 493 513 POTENTIAL.
FT MOD_RES 28 28 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 594 594 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 674 AA; 71626 MW; 67A75A8F76E42FA2 CRC64;

Query Match 2.28; Score 11; DB 1; Length 674;
Best Local Similarity 100.0%; Pred. No. 0.042; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

OY 293 GAAAGGAGAA 303
DB 499 GAAAGGAGAA 509

RESULT 3
FSH_DROME STANDARD; PRT; 2038 AA.
ID FSH_DROME
AC P13709; P13710; (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Female sterile homeotic protein (Fragile-chorion membrane protein).
GN FSH OR FSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89276730; PubMed=2567251;
RA Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
RT "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
RT apparent membrane proteins.";
RL Dev. Biol. 134:246-257(1989)
CC -1- FUNCTION: REQUIRED MATERIALLY FOR PROPER EXPRESSION OF OTHER
CC HOMEOTIC GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
CC -1- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -1- SIMILARITY: CONTAINS 1 ET DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23221; AAA28540.1; -
CC EMBL; M23222; AAA28541.1; ALT TERM.
CC EMBL; M15762; AAA70424.1; -
CC EMBL; M15763; AAA70423.1; -
CC EMBL; M15764; AAA70422.1; -
CC PIR; A43742; A43742.
CC HSSP; Q92831; 1B91.
CC FlyBase; FBgn0004656; fs(1)h.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 2.
CC PROSITE; PS00633; BROMODOMAIN_1; 2.
CC PROSITE; PS00144; BROMODOMAIN_2; 2.
CC Developmental protein; Bromodomain; Transmembrane; Repeat.
KW DOMAIN 51 123 BROMODOMAIN 1.
FT DOMAIN 945 1106 ET DOMAIN.

```

```

FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 750 770 POTENTIAL.
FT TRANSMEM 790 810 POTENTIAL.
FT TRANSMEM 816 830 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 1731 1751 POTENTIAL.
FT TRANSMEM 1939 1959 POTENTIAL.
FT VARIANT 909 909 G -> A.
FT VARIANT 1022 1022 H -> RPKY.
SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 2.28; Score 11; DB 1; Length 2038;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAGAGGAGAA 304
DB 343 AAGAGGAGAA 353

RESULT 4
RS21_RHIL0 STANDARD; PRT; 88 AA.
ID RS21_RHIL0
AC O98GY4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S21.
GN RPSU OR MSR3117.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OC NCBI_Taxid=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099; PubMed=11214968;
RX Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RX Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RX Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
RX Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RX Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003001; BAB50082.1; -
CC InterPro; IPR001911; Ribosomal_S21.
CC Pfam; PF01163; Ribosomal_S21.
CC PRINTS; PR00976; RIBOSOMAL_S21.
CC TRIFRAMS; TIGR00030; S21P; 1.
CC PROSITE; PS01181; RIBOSOMAL_S21; FALSE_NEG.
CC Ribosomal protein; Complete proteome.
KW SEQUENCE 88 AA; 10045 MW; 2C8536956D632157 CRC64;

Query Match 1.88; Score 9; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGAGGAA 304
DB 76 AAGAGGAA 84

```

```

983 6 1.2 364 1 A3_VIGUN Q41706 vigna ungu
984 6 1.2 364 1 FLGI_ZYMO Q2519 zymomonas m
985 6 1.2 364 1 SYM_MERTH Q26352 methanobact
986 6 1.2 365 1 MN16_HUMAN Q9ubv4 homo sapien
987 6 1.2 366 1 ALF2_CAEEL P46563 caenorhabdi
988 6 1.2 366 1 METK_PEA P49612 pisum sativ
989 6 1.2 367 1 ELV3_HUMAN Q14576 homo sapien
990 6 1.2 367 1 ELV3_MOUSE Q60900 mus musculu
991 6 1.2 367 1 NADA_BACHD Q9k433 bacillus ha
992 6 1.2 367 1 MECA_SALRT Q82386 salmoneilla
993 6 1.2 367 1 MECA_SALTY Q91677 salmoneilla
994 6 1.2 368 1 GAL7_HUMAN Q60755 homo sapien
995 6 1.2 368 1 H181_RHIME Q92mg0 rhizobium m
996 6 1.2 368 1 HIS8_AGR5 Q8ubw3 agrobacteri
997 6 1.2 368 1 HIS8_BRUME Q8y1k3 bruceella me
998 6 1.2 368 1 HIS8_BUCAT P57202 buchnera ap
999 6 1.2 368 1 YTL_MOUSE Q62481 mus musculu
1000 6 1.2 369 1 FLGI_PSEAE Q914p5 pseudomonas

```

## ALIGNMENTS

## RESULT 1

RL7\_RHIME

ID RL7\_RHIME STANDARD: PRT: 126 AA.

AC 0920H8:

DT 15-JUN-2002 (rel. 41, Created)

DT 15-JUN-2002 (rel. 41, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE 50S ribosomal protein L7/L12.

GN RPL7 OR R01347 OR SMC01318.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

CC Rhizobiaceae; Sinorhizobium.

OX NCBI\_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Bailly-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Bolstad P., Becker A., Boutry M., Cadieu E., Diano S., Gloux S.,

RA Godie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,

RA Pohl T., Portetalle D., Puhler A., Purnelle B., Ramsperger U.,

RA Renard C., Thibault P., Vandenbol M., Weidner S., Gallbert F.,

RT Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021.

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC - FUNCTION: Seems to be the binding site for several of the factors

CC involved in protein synthesis and appears to be essential for

CC accurate translation (By similarity).

CC - SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AL591787; CAC45926.1; -

CC InterPro: IPR000206; Ribosomal\_L12.

CC Pfam: PF00542; Ribosomal\_L12; 1.

CC ProDom: PD001326; Ribosomal\_L12; 1.

CC TIGRfams: TIGR00855; L12; 1.

CC Ribosomal protein; Complete proteome.

CC SEQUENCE 126 AA; 12835 MW; 71D2BF16FC487F88 CRC64;

SQ

Query Match 2.2%; Score 11; DB 1; Length 126;

Best local similarity 100.0%; Pred. No. 0.0096;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 295 AAAGAGAGAAA 305
DB 41 AAAGAGAGAAA 51

```

## RESULT 2

PTGA\_CORGL

ID PTGA\_CORGL STANDARD: PRT: 674 AA.

AC 045298:

DT 01-NOV-1997 (rel. 35, Created)

DT 01-NOV-1997 (rel. 35, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE PTS system, glucose-specific IYABC component (EIYABC-GLC) (Glucose-

DE permease IYABC component) (Phosphotransferase enzyme II, ABC

DE component) (EC 2.7.1.69) (EI-GLC/EIY-GLC).

GN PTSG.

OS Corynebacterium glutamicum (Brevibacterium flavum).

OG Plasmid pBSBQ2.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;

CC Corynebacterium.

OX NCBI\_TaxID=1718;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13869;

RA Yoon K.-H.;

RT Cloning and nucleotide sequence of enzyme II of Brevibacterium

RT lactofermentum phosphotransferase system.

RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.

CC - FUNCTION: THIS IS A COMPONENT OF THE PHENOLPYRUVATE-DEPENDENT

CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE

CC -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE

CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY

CC PHOSPHORYLATION SITE (THE DONOR IS PHOSHO-HPR); IIA TRANSFERS ITS

CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO

CC THE SUGAR.

CC - CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein

CC histidine + sugar phosphate.

CC - SUBCELLULAR LOCATION: Integral membrane protein.

CC - SIMILARITY: CONTAINS 1 PTS EIIA DOMAIN.

CC - SIMILARITY: CONTAINS 1 PTS EIIC DOMAIN.

CC - SIMILARITY: CONTAINS 1 PTS EIIC DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: L18875; AAA22992.1; -

CC HSSP: P08837; IGIC.

CC InterPro: IPR001127; PTS\_EIIA.

CC InterPro: IPR001996; PTS\_EIIB.

CC InterPro: IPR003352; PTS\_EIIC.

CC Pfam: PF00358; PTS\_EIIA.1; 1.

CC Pfam: PF00367; PTS\_EIIB.1; 1.

CC Pfam: PF02378; PTS\_EIIC.1; 1.

CC ProDom: PD001476; PTS\_EIIB.1; 1.

CC ProDom: PD002243; PTS\_EIIB.1; 1.

CC TIGRfams: TIGR00830; PTBA.1; 1.

CC PROSITE: PS00371; PTS\_EIIA.1; 1.

CC PROSITE: PS01035; PTS\_EIIB-CYS.1.

CC Phosphotransferase system; Sugar transport; Transferase;

CC Phosphorylation; Transmembrane; Plasmid.

CC DOMAIN 1 43

CC DOMAIN 2 674

CC DOMAIN 3 146

CC DOMAIN 4 182

CC DOMAIN 5 213

CC DOMAIN 6 245

CC TRANSMEM 126

CC TRANSMEM 162

CC TRANSMEM 193

CC TRANSMEM 225

CC TRANSMEM 245

CC TRANSMEM 245

CC TRANSMEM 245

CC TRANSMEM 245

CC TRANSMEM 245



837	6	1.2	315	1	COXX_BRUME	Q8f96 bruceella me	910	6	1.2	339	1	G3P_BRUMA	P48812 brugia mala
838	6	1.2	315	1	HYFC_ECOLI	P77858 escherichia	911	6	1.2	339	1	G3P_ONCVO	O01360 onchocerca
839	6	1.2	315	1	SHX2_HUMAN	O15370 homo sapien	912	6	1.2	339	1	HEMZ_RHIME	O92m52 rhizobium m
840	6	1.2	315	1	Y060_BUCAT	P57168 buchneria ap	913	6	1.2	339	1	PLSX_CLOPE	O8xj96 clostridium
841	6	1.2	317	1	Y06L_ECOLI	P29744 escherichia	914	6	1.2	339	1	PYRD_HAEIN	P45477 haemophilus
842	6	1.2	317	1	MTX1_HUMAN	O13505 homo sapien	915	6	1.2	339	1	PYRD_PASMU	P57858 pasteurella
843	6	1.2	317	1	MTX1_MOUSE	P47802 mus musculu	916	6	1.2	339	1	RIBD_MYCTU	P71677 m riboflavi
844	6	1.2	317	1	RLAO_HUMAN	P05388 homo sapien	917	6	1.2	341	1	ARGC_METJA	O58496 methanococ
845	6	1.2	317	1	RLAO_MOUSE	P14669 mus musculu	918	6	1.2	341	1	JUND_MOUSE	P15066 mus musculu
846	6	1.2	317	1	RLAO_RAT	P19495 ratulus norv	919	6	1.2	341	1	HUPK_AZCVI	P52909 rattus norv
847	6	1.2	317	1	SOX2_HUMAN	P48431 homo sapien	920	6	1.2	342	1	TRM1_THEVO	O97ir2 thermoplas
848	6	1.2	317	1	ADT2_YEAST	O03241 bovine rota	921	6	1.2	342	1	GPDA_RHILLO	P58142 rhizobium 1
849	6	1.2	318	1	ETFA_MYCLE	P18239 saccharomyc	922	6	1.2	343	1	HRCA_MYCGE	P47447 mycoplasma
850	6	1.2	318	1	ETFA_MYCTU	O53275 mycobacteri	923	6	1.2	344	1	CD2_MOUSE	P08920 mus musculu
851	6	1.2	318	1	IFR_MEDSA	P52575 medicago sa	924	6	1.2	344	1	CEA6_HUMAN	P40199 homo sapien
852	6	1.2	319	1	CCSA_OENHO	O9m112 oenothera h	925	6	1.2	344	1	COMC_METJA	O58820 methanococ
853	6	1.2	319	1	PSTC_MYCLE	P17110 ceratitlis c	926	6	1.2	344	1	MPRR_STRLI	P43161 streptomyc
854	6	1.2	320	1	CH36_CERCA	P56164 phalaris aq	927	6	1.2	346	1	ULB6_HCMVA	P16833 human cytom
855	6	1.2	320	1	MP51_PHANQ	O13630 homo sapien	928	6	1.2	346	1	PUR5_BACSU	P23715 bacillus su
856	6	1.2	320	1	SOX2_SHEP	P25954 bacillus su	929	6	1.2	346	1	VC54_SYNY3	P12043 synochocyst
857	6	1.2	321	1	FCL_HUMAN	P1652 ratulus norv	930	6	1.2	347	1	ARGC_BACSU	P74078 synochocyst
858	6	1.2	321	1	CMGB_BACSU	P58000 erwinaa her	931	6	1.2	347	1	PUR5_BACSU	O52956 mycobacteri
859	6	1.2	323	1	PE2R_RAT	O86344 mycobacteri	932	6	1.2	347	1	ADH2_YEAST	P00331 saccharomyc
860	6	1.2	323	1	TKRA_ERHME	P23617 fusarium so	933	6	1.2	347	1	ARGC_CORGL	O59279 corynebacte
861	6	1.2	324	1	THI4_FUSSH	P21160 mycobacteri	934	6	1.2	347	1	ID12_LACIA	O9c1f5 lactococcus
862	6	1.2	325	1	A85B_MYCNA	O95886 homo sapien	935	6	1.2	347	1	JUND_HUMAN	P17275 homo sapien
863	6	1.2	325	1	DLP3_HUMAN	P44551 bacillus su	936	6	1.2	347	1	PUR7_HALNL	P17353 homo sapien
864	6	1.2	325	1	ETFA_BACSU	P3651 streptomyc	937	6	1.2	347	1	TOIA_PSEAE	Q9hnuv halobacteri
865	6	1.2	326	1	BLAA_STRCI	P94284 borelia bu	938	6	1.2	347	1	HOXV_AZOVI	P50600 pseudomonas
866	6	1.2	326	1	TRXB_BORBU	P31317 schizosacch	939	6	1.2	347	1	ASG2_HAEIN	O91025 gallus gall
867	6	1.2	326	1	VS09_ROTBR	O05826 gallus gall	940	6	1.2	348	1	DRK3_MOUSE	P40597 azotobacter
868	6	1.2	326	1	ORC_SCHPO	O91127 streptomyc	941	6	1.2	349	1	LEP_HAEIN	P43843 haemophilus
869	6	1.2	327	1	CEBB_CHICK	O15077 homo sapien	942	6	1.2	349	1	HMRO_DROME	P41394 leptospira
870	6	1.2	328	1	MPR2_STRCO	O63371 ratulus norv	943	6	1.2	350	1	OPSR_CONCO	O94un9 haemophilus
871	6	1.2	328	1	P2Y6_HUMAN	O34436 bacillus su	944	6	1.2	350	1	HRCA_MYCPN	P44544 haemophilus
872	6	1.2	328	1	PIT_BACSU	O09461 caenorhabdi	945	6	1.2	351	1	TAL_CANAL	P10181 drosophila
873	6	1.2	328	1	YO51_CAEEL	P17819 kluyveromyc	946	6	1.2	351	1	TAL_NEIMB	O13227 conger cong
874	6	1.2	328	1	G3P_KIUIA	P54097 gallus gall	947	6	1.2	351	1	VORB_METTH	O9hvx0 pseudomonas
875	6	1.2	329	1	TECH_CHICK	O06947 mycobacteri	948	6	1.2	351	1	EGSA_THEVO	P75351 mycoplasma
876	6	1.2	329	1	A85B_MYCAY	O49575 mycobacteri	949	6	1.2	351	1	OPSD_CONCO	O94091 candida alb
877	6	1.2	330	1	A85B_MYCIT	O43488 homo sapien	950	6	1.2	351	1	OPSD_ZOSOP	O9tix7 cytidium c
878	6	1.2	330	1	AR72_HUMAN	P06877 escherichia	951	6	1.2	352	1	CHLI_CVACA	O27089 methanobact
879	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	952	6	1.2	352	1	OPSD_POMTI	O92009 mus musculu
880	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	953	6	1.2	352	1	OPSD_ZOSOP	O9h116 thermoplas
881	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	954	6	1.2	352	1	EGSA_THEVO	O9y422 goblus nige
882	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	955	6	1.2	352	1	OPSD_ZOSOP	P35403 pomatoschis
883	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	956	6	1.2	352	1	OPSD_ZOSOP	O9y8y9 zosterisess
884	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	957	6	1.2	352	1	OPSD_ZOSOP	O9tix7 cytidium c
885	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	958	6	1.2	352	1	OPSD_ZOSOP	O27089 methanobact
886	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	959	6	1.2	352	1	OPSD_ZOSOP	O92009 mus musculu
887	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	960	6	1.2	352	1	OPSD_ZOSOP	O9h116 thermoplas
888	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	961	6	1.2	352	1	OPSD_ZOSOP	O9y422 goblus nige
889	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	962	6	1.2	352	1	OPSD_ZOSOP	P35403 pomatoschis
890	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	963	6	1.2	352	1	OPSD_ZOSOP	O9y8y9 zosterisess
891	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	964	6	1.2	352	1	OPSD_ZOSOP	O9tix7 cytidium c
892	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	965	6	1.2	352	1	OPSD_ZOSOP	O27089 methanobact
893	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	966	6	1.2	352	1	OPSD_ZOSOP	O92009 mus musculu
894	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	967	6	1.2	352	1	OPSD_ZOSOP	O9h116 thermoplas
895	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	968	6	1.2	352	1	OPSD_ZOSOP	O9y422 goblus nige
896	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	969	6	1.2	352	1	OPSD_ZOSOP	P35403 pomatoschis
897	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	970	6	1.2	352	1	OPSD_ZOSOP	O9y8y9 zosterisess
898	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	971	6	1.2	352	1	OPSD_ZOSOP	O9tix7 cytidium c
899	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	972	6	1.2	352	1	OPSD_ZOSOP	O27089 methanobact
900	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	973	6	1.2	352	1	OPSD_ZOSOP	O92009 mus musculu
901	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	974	6	1.2	352	1	OPSD_ZOSOP	O9h116 thermoplas
902	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	975	6	1.2	352	1	OPSD_ZOSOP	O9y422 goblus nige
903	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	976	6	1.2	352	1	OPSD_ZOSOP	P35403 pomatoschis
904	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	977	6	1.2	352	1	OPSD_ZOSOP	O9y8y9 zosterisess
905	6	1.2	330	1	G3P1_ECOLI	O01558 leishmania	978	6	1.2	352	1	OPSD_ZOSOP	O9tix7 cytidium c
906	6	1.2	330	1	G3P1_ECOLI	O41629 triticum ae	979	6	1.2	352	1	OPSD_ZOSOP	O27089 methanobact
907	6	1.2	330	1	G3P1_ECOLI	O41630 triticum ae	980	6	1.2	352	1	OPSD_ZOSOP	O92009 mus musculu
908	6	1.2	330	1	G3P1_ECOLI	P15286 escherichia	981	6	1.2	352	1	OPSD_ZOSOP	O9h116 thermoplas
909	6	1.2	330	1	G3P1_ECOLI	P24165 salmonella	982	6	1.2	352	1	OPSD_ZOSOP	O9y422 goblus nige



691	6	1.2	272	1	YW26_SULSO	Q97u03 sulfolobus	764	6	1.2	298	1	ADT2_HUMAN	P05141 homo sapien
692	6	1.2	273	1	NK2B_HUMAN	O95096 homo sapien	765	6	1.2	298	1	ADT2_MOUSE	P51881 mus musculu
693	6	1.2	273	1	PRIO_CHICK	P27177 gallus gall	766	6	1.2	298	1	ADT2_RAT	O09073 rattus norv
694	6	1.2	274	1	YOHM_ECOLI	P76425 escherichia	767	6	1.2	298	1	ADT3_BOVIN	P32007 bos taurus
695	6	1.2	275	1	HXA5_HETER	Q91a23 heterodontu	768	6	1.2	298	1	ADT3_HUMAN	P12236 homo sapien
696	6	1.2	275	1	LEP4_BURPS	Q9zft0 burkholderi	769	6	1.2	298	1	COXX_PARDE	P08301 paracoccus
697	6	1.2	275	1	MYO2_ONCMY	O91206 oncohynechu	770	6	1.2	298	1	ICIA_VIBCH	Q9kun3 vibrio chol
698	6	1.2	276	1	OCCF_AGRPU	P35121 agrobacteri	771	6	1.2	298	1	LIPA_BACSU	O92129 bacillus su
699	6	1.2	276	1	SW21_HUMAN	O9z651 homo sapien	772	6	1.2	298	1	MTRE_MERKA	O94606 methanopyru
700	6	1.2	277	1	IMD_STRCN	O33611 streptomyc	773	6	1.2	298	1	OXIR_STRAT	Q03326 streptomyc
701	6	1.2	277	1	Y4OR_RHISN	P55603 rhizobium s	774	6	1.2	298	1	PRIB_LACPL	P77867 lactobacill
702	6	1.2	278	1	CHIS_STRSO	P33665 streptomyc	775	6	1.2	299	1	BAH_STRMY	Q01109 streptomyc
703	6	1.2	278	1	T2D7_DROME	Q27272 drosophila	776	6	1.2	299	1	HEVI_MOUSE	Q9wv93 mus musculu
704	6	1.2	279	1	LEP4_ERWCA	P31712 erwilia car	777	6	1.2	299	1	HM30_CAEBL	Q22909 caenorhadi
705	6	1.2	279	1	Y284_CHVPI	P54966 paramectium	778	6	1.2	299	1	ISPE_STRCO	Q9k3r6 streptomyc
706	6	1.2	280	1	AEOK_AERPE	O9yek6 aetopyrium p	779	6	1.2	299	1	MP51_MYCTU	Q48922 mycobacteri
707	6	1.2	280	1	LE76_BRANA	P13934 brassica na	780	6	1.2	299	1	PRTB_STRGR	P00777 streptomyc
708	6	1.2	280	1	NADG_AERPE	O9yval1 aetopyrium p	781	6	1.2	299	1	YHJC_ECOLI	P37641 escherichia
709	6	1.2	280	1	PRTA_ASPNG	P24665 deinooccus	782	6	1.2	300	1	ARGB_BRUME	O8yda5 bruceella me
710	6	1.2	282	1	DRRB_STRPE	P32011 streptomyc	783	6	1.2	301	1	ADT_AMOGA	Q27238 anopheltes g
711	6	1.2	283	1	MAUN_PARDE	O51660 paracoccus	784	6	1.2	301	1	HYDD_CHICK	P24344 gallus gall
712	6	1.2	283	1	YK26_YEAST	P36139 saccharomyc	785	6	1.2	301	1	OTCA_VIBS2	P66172 vibrio chol
713	6	1.2	283	1	COXX_RHOSH	P56938 rhodobacter	786	6	1.2	301	1	Y186_MYCPN	P75265 mycoplasma
714	6	1.2	284	1	KIME_ARCFU	O27995 archaeoglob	787	6	1.2	302	1	HHPX_AOUAE	O67798 aquilex leo
715	6	1.2	284	1	MP5B_PHLPR	O40963 phleum prat	788	6	1.2	302	1	RRAW_ANASP	O8yrt9 anabaena sp
716	6	1.2	284	1	PMXA_HUMAN	O14813 homo sapien	789	6	1.2	302	1	RECR_CHRVI	P25544 chromalium
717	6	1.2	284	1	BLA1_ENTCL	P18251 enterobacte	790	6	1.2	303	1	E134_MAIZE	O9zt66 zea mays l m
718	6	1.2	286	1	ILVD_CLOPA	P31959 clostridium	791	6	1.2	303	1	LIPA_BACBD	O9y7c9 bacillus ha
719	6	1.2	286	1	MP5A_PHLPR	O40962 phleum prat	792	6	1.2	303	1	YNA6_YEAST	P41318 saccharomyc
720	6	1.2	286	1	RER2_YEAST	P35196 saccharomyc	793	6	1.2	304	1	HEV1_CANPA	Q9fsz2 canis famli
721	6	1.2	286	1	Y320_MYCGE	P47562 mycoplasma	794	6	1.2	304	1	HEV1_HUMAN	O9y513 homo sapien
722	6	1.2	286	1	Y144_KIUUA	P31034 kluyveromyc	795	6	1.2	304	1	MRR_ECOLI	P24202 escherichia
723	6	1.2	286	1	YMB1_CAEBL	P50093 caenorhadi	796	6	1.2	305	1	AOV1_YEAST	P35386 saccharomyc
724	6	1.2	287	1	ATPG_ECOLI	P00837 escherichia	797	6	1.2	305	1	BLAC_STRLA	P35393 streptomyc
725	6	1.2	287	1	DNAJ_SYNPN	P50026 synecococc	798	6	1.2	305	1	Y452_BUCAT	P57527 buchnera ap
726	6	1.2	287	1	RL3_MYCPN	P75580 mycoplasma	799	6	1.2	306	1	OPBC_BACSU	Q45462 bacillus su
727	6	1.2	287	1	YGE7_YEAST	P53171 saccharomyc	800	6	1.2	307	1	ADT3_YEAST	O9y513 homo sapien
728	6	1.2	288	1	ISPE_CHLTR	O84810 chlamydia t	801	6	1.2	307	1	EMFA_PAREE	P38974 paracoccus
729	6	1.2	288	1	Y077_HAETN	P43935 haemophilus	802	6	1.2	307	1	MP93_PORAR	P27286 poa pratens
730	6	1.2	288	1	FMT_MYCPU	O98194 mycoplasma	803	6	1.2	307	1	PM34_MOUSE	O70579 mus musculu
731	6	1.2	289	1	Y465_TREPA	O83478 treponema p	804	6	1.2	307	1	SCRK_KLEPN	P26420 klebsiella
732	6	1.2	290	1	Y647_HAETN	P45293 haemophilus	805	6	1.2	307	1	TRUB_PASMU	O9cmq7 pasteurella
733	6	1.2	291	1	BIP2_TOBAC	O03682 nicotiana t	806	6	1.2	307	1	YEAT_ECOLI	P76250 escherichia
734	6	1.2	292	1	GALU_MYCGE	P47691 mycobactera	807	6	1.2	307	1	YOJK_BACSU	P54548 bacillus su
735	6	1.2	292	1	HYRB_RHOCA	P42507 rhodobacter	808	6	1.2	307	1	ADT_CHLRE	P27080 chlamydomon
736	6	1.2	293	1	BIP8_TOBAC	O03686 nicotiana t	809	6	1.2	308	1	MACS_MOUSE	P26645 mus musculu
737	6	1.2	293	1	HMX1_MOUSE	P13297 mus musculu	810	6	1.2	308	1	MACS_RAT	P30009 rattus norv
738	6	1.2	293	1	PH80_YEAST	P20052 saccharomyc	811	6	1.2	308	1	VPT_HAETN	P44227 haemophilus
739	6	1.2	293	1	ARGB_AGRF5	O8u1b7 agrobacteri	812	6	1.2	309	1	AGAR_STRCO	O78823 streptomyc
740	6	1.2	294	1	G3P_ESCFE	P24746 escherichia	813	6	1.2	309	1	TRUB_YEREN	O34273 yersinia en
741	6	1.2	294	1	MMSB_MYCTU	O53814 mycobacteri	814	6	1.2	310	1	CCW7_YEAST	P47000 saccharomyc
742	6	1.2	294	1	MP53_PHAAP	P56166 phalaris ag	815	6	1.2	310	1	TRUB_CAUCR	P38063 caulobacter
743	6	1.2	294	1	POL_SMSAV	O92359 simian sarc	816	6	1.2	310	1	Y063_SYN3	P74068 synecocyst
744	6	1.2	294	1	ARGB_RHIME	O92390 rhizobium m	817	6	1.2	311	1	TYTP_BOVIN	P79110 bos taurus
745	6	1.2	295	1	ISPE_VIBCH	O9Kq23 vibrio chol	818	6	1.2	311	1	TYTP_HUMAN	P53007 homo sapien
746	6	1.2	295	1	MEDE_YEAST	P38782 saccharomyc	819	6	1.2	312	1	CC2D_ANTMA	O43142 haemophilus
747	6	1.2	295	1	MTRE_METTH	O27231 methanobact	820	6	1.2	312	1	TRUB_HAETN	P45147 streptomyc
748	6	1.2	295	1	MTRE_METTH	P80186 methanobact	821	6	1.2	312	1	TRUB_VIBCH	O9kuv8 vibrio chol
749	6	1.2	296	1	CRTB_PANAN	P21683 pantoea ana	822	6	1.2	312	1	ADT_NEUCR	P02723 neurospora
750	6	1.2	296	1	DAPE_MYCLE	O46814 mycobacteri	823	6	1.2	313	1	CAS1_MOUSE	P19228 mus musculu
751	6	1.2	296	1	FCG2_BOVIN	O28110 bos taurus	824	6	1.2	313	1	CIW6_HUMAN	O9y257 homo sapien
752	6	1.2	296	1	HO_CHICK	P14791 gallus gall	825	6	1.2	313	1	KREL_YEAST	P17260 saccharomyc
753	6	1.2	296	1	SAPR_STRPU	O54262 xanthomonas	826	6	1.2	313	1	Y085_CAEBL	P08391 escherichia
754	6	1.2	296	1	Y300_METJA	O57748 methanococc	827	6	1.2	313	1	FMT_ECOS7	P36621 caenorhadi
755	6	1.2	296	1	Y799_METJA	O58209 methanococc	828	6	1.2	314	1	FMT_ECOS7	O8x8f1 caenorhadi
756	6	1.2	296	1	ADT1_BOVIN	O62722 bos taurus	829	6	1.2	314	1	FMT_PSEAE	P23882 escherichia
757	6	1.2	297	1	ADT1_BOVIN	O26365 dirosophila	830	6	1.2	314	1	FMT_SALTY	O85732 pseudomonas
758	6	1.2	297	1	ADT1_BOVIN	O98476 rhizobium l	831	6	1.2	314	1	PMXB_HUMAN	O8z1x0 salmonella
759	6	1.2	297	1	ADT1_HUMAN	P12235 homo sapien	832	6	1.2	314	1	PMXB_MOUSE	O99453 homo sapien
760	6	1.2	298	1	ADT1_HUMAN	P48962 mus musculu	833	6	1.2	314	1	TRUB_ECOLI	O35690 mus musculu
761	6	1.2	298	1	ADT1_MOUSE	Q05962 rattus norv	834	6	1.2	314	1		
762	6	1.2	298	1	ADT1_MOUSE		835	6	1.2	314	1		
763	6	1.2	298	1	ADT1_MOUSE		836	6	1.2	314	1		

us-09-889-314-2.rsp

Tue Jan 28 10:21:48 2003

545	6	1.2	228	1	V33P_ADE02	P24939 human adeno	618	6	1.2	254	1	RACD-DICD1	P34150 dictyosteli
546	6	1.2	229	1	TDX1_BR0M4	P48822 brugia mala	619	6	1.2	255	1	TNF9_HUMAN	P41273 homo sapien
547	6	1.2	229	1	V33P_ADE05	P24940 human adeno	620	6	1.2	255	1	PRI0_CAMDR	P79141 camelus dro
548	6	1.2	231	1	PCAI_PSEPU	P01103 pseudomonas	621	6	1.2	255	1	PRI0_CAMDR	P46501 canis famli
549	6	1.2	231	1	YOHK_HAEIN	P45146 haemophilus	622	6	1.2	255	1	YPOH_RHOVI	P20467 rhodospendo
550	6	1.2	232	1	PRI0_ATEGE	P40246 ateles geof	623	6	1.2	256	1	ECHE_MYCTU	O53211 mycobacteri
551	6	1.2	233	1	HEX8_ADE40	P11821 human adeno	624	6	1.2	256	1	L301_BOMMO	O00802 bromyces fa
552	6	1.2	233	1	HEX8_ADE41	P11822 human adeno	625	6	1.2	256	1	MDTH_UROFA	P52113 capra hircu
553	6	1.2	233	1	SODM_YEAST	P00447 saccharomyc	626	6	1.2	256	1	PRI0_CAPIH	P79142 cervus elap
554	6	1.2	233	1	SDX20_HUMAN	P060248 homo sapien	627	6	1.2	256	1	PRI0_FELCA	P15082 odococleus
555	6	1.2	234	1	TFZD_ENTHI	P21653 entamoeba n	628	6	1.2	256	1	PRI0_OPOHE	P23907 ovis aures
556	6	1.2	234	1	TFZD_ENTHI	P07136 alcaicogenes	629	6	1.2	256	1	PRI0_SHEEP	P47852 odococleus
557	6	1.2	235	1	FLGH_RHILLO	P08439 methanococc	630	6	1.2	256	1	PRP2_BOVIN	P40243 tragelaphus
558	6	1.2	235	1	YA99_METVA	P049730 arabisidopsis	631	6	1.2	257	1	PRP2_BOVIN	P35174 musstela put
559	6	1.2	236	1	MRAY_ARATH	O95197 homo sapien	632	6	1.2	257	1	FAB1_ECOLI	P40244 mustela vis
560	6	1.2	236	1	RTN3_HUMAN	O95197 homo sapien	633	6	1.2	257	1	PRI0_MUSPI	P49927 sus scrofa
561	6	1.2	237	1	PYRE_FUSNM	O81983 fusobacteri	634	6	1.2	257	1	PRI0_PIG	P47397 mycoplasma
562	6	1.2	237	1	MEUM_HUMAN	P17677 homo sapien	635	6	1.2	257	1	RL3_MYCGE	P15082 escherichia
563	6	1.2	238	1	PRI0_CERAT	O95145 ceratococcus	636	6	1.2	257	1	RL3_MYCGE	O9517 lactococcus
564	6	1.2	238	1	PRI0_THERG	O95270 theropithec	637	6	1.2	257	1	SLR1_ECOLI	P38872 bos taurus
565	6	1.2	239	1	PRI0_AOTTR	P40245 aotus trivi	638	6	1.2	258	1	RNH2_LACIA	P31780 trichosturus
566	6	1.2	240	1	YDGB_ECOLI	P52109 escherichia	639	6	1.2	258	1	UPKA_BOVIN	P53269 saccharomyc
567	6	1.2	241	1	MLTE_ECOLI	P76009 escherichia	640	6	1.2	259	1	PRI0_TRIUV	O53774 mycobacteri
568	6	1.2	241	1	PRI0_CALMO	P40248 callicebus	641	6	1.2	259	1	YG31_YEAST	P53595 methylobact
569	6	1.2	241	1	PRI0_MANSP	O40255 mandrillus	642	6	1.2	260	1	CE30_MYCTU	P40258 salmistr sci
570	6	1.2	243	1	AX17_ARATH	O38825 arabidopsis	643	6	1.2	260	1	MTXB_METEX	O95911 natrialba m
571	6	1.2	243	1	YN06_CAEEL	O03615 caenorhabdi	644	6	1.2	261	1	PRI0_SAISS	P39698 bacterioph
572	6	1.2	243	1	PRI0_CERBE	P40250 ceratococcus	645	6	1.2	261	1	FLA4_NATMA	O93912 tetrastoma s
573	6	1.2	245	1	TRT3_ECOLI	P13980 escherichia	646	6	1.2	261	1	VBET_LAMB	P42452 synechococ
574	6	1.2	245	1	Y355_HAEIN	O57306 haemophilus	647	6	1.2	262	1	Y558_METVA	P95156 mycobacteri
575	6	1.2	246	1	PMG1_YEAST	OBY124 brucella me	648	6	1.2	262	1	IFZ2_METTH	P10279 bos taurus
576	6	1.2	246	1	PRI0_CERMO	P00950 saccharomyc	649	6	1.2	262	1	YNT2_MYCTU	P40242 tregelaphus
577	6	1.2	246	1	PRI0_CERPA	O95174 ceratococcus	650	6	1.2	263	1	DAPE_RALSO	P27380 bacterioph
578	6	1.2	246	1	PRI0_CERPA	O95174 ceratococcus	651	6	1.2	263	1	DAPE_RALSO	P50644 homo sapien
579	6	1.2	246	1	VMRY_LAMB	O95174 ceratococcus	652	6	1.2	264	1	HEM4_STYF7	O94944 mus musculu
580	6	1.2	246	1	SN61_CAEEL	P03733 bacterioph	653	6	1.2	264	1	MODB_MYCTU	P47864 ratus norv
581	6	1.2	247	1	UCR1_SPTOL	O76735 spinacia ol	654	6	1.2	264	1	PRI0_BOVIN	O05775 saccharomyc
582	6	1.2	247	1	YCP4_YEAST	P25349 saccharomyc	655	6	1.2	264	1	PRP1_BOVIN	O05775 saccharomyc
583	6	1.2	248	1	DSBG_ECO57	P58320 escherichia	656	6	1.2	264	1	VP07_BPRPD	O02184 synechococ
584	6	1.2	248	1	DSBG_ECO57	P72302 escherichia	657	6	1.2	264	1	VP07_BPRPD	P27685 dictyosteli
585	6	1.2	249	1	CRB1_RAT	P02523 ratus norv	658	6	1.2	265	1	ADP5_HUMAN	O9392 drosophila
586	6	1.2	249	1	KTXG_KL01A	P09807 kluyveromyc	659	6	1.2	265	1	ADP5_MOUSE	P51373 porphyra pu
587	6	1.2	249	1	PMGY_MYCTU	O11140 mycobacteri	660	6	1.2	265	1	ADP5_MOUSE	P34444 caenorhabdi
588	6	1.2	249	1	PMGY_MYCTU	P66190 xanthopacta	661	6	1.2	265	1	ADP5_MOUSE	O92420 caulobacter
589	6	1.2	249	1	PMGY_MYCTU	P20682 podospora a	662	6	1.2	265	1	ADP5_MOUSE	P32645 saccharomyc
590	6	1.2	250	1	FCG3_BOVIN	P79107 dos taurus	663	6	1.2	265	1	ADP5_MOUSE	O56028 methanosarc
591	6	1.2	250	1	FCG3_BOVIN	P30793 homo sapien	664	6	1.2	265	1	ADP5_MOUSE	P20026 hordeum vul
592	6	1.2	251	1	GCRI_HUMAN	P39447 homo sapien	665	6	1.2	265	1	ADP5_MOUSE	O94944 mus musculu
593	6	1.2	251	1	GCRI_HUMAN	P33890 saccharomyc	666	6	1.2	265	1	ADP5_MOUSE	O05775 saccharomyc
594	6	1.2	251	1	TIR2_YEAST	P50918 lactococcus	667	6	1.2	265	1	ADP5_MOUSE	O05775 saccharomyc
595	6	1.2	252	1	TIR2_YEAST	P51446 ateles panl	668	6	1.2	265	1	ADP5_MOUSE	O05775 saccharomyc
596	6	1.2	252	1	PRI0_ATEPA	P40247 callithrix	669	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
597	6	1.2	252	1	PRI0_CALJA	P40247 callithrix	670	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
598	6	1.2	252	1	PRI0_CEBAP	O95211 cryocollagus	671	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
599	6	1.2	252	1	PRI0_RABIT	O95211 cryocollagus	672	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
600	6	1.2	253	1	EXBB_XANCP	O34260 xanthomomas	673	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
601	6	1.2	253	1	PRI0_GOLGU	P40251 colobus que	674	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
602	6	1.2	253	1	PRI0_GOLGU	P40252 gorilla gor	675	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
603	6	1.2	253	1	PRI0_HOGAN	P04156 homo sapien	676	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
604	6	1.2	253	1	PRI0_MACEA	P40254 macaca fasc	677	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
605	6	1.2	253	1	PRI0_PANTR	P40253 pan troglod	678	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
606	6	1.2	253	1	PRI0_PONPY	P40256 pongo pygma	679	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
607	6	1.2	253	1	PRI0_PREFR	P40257 presbytis st	680	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
608	6	1.2	253	1	TPIS_BACST	P00453 bacillus st	681	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
609	6	1.2	254	1	TPIS_BACST	P00453 bacillus st	682	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
610	6	1.2	254	1	HM36_CAEEL	P40067 haemophilus	683	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
611	6	1.2	254	1	HM36_CAEEL	O93352 baomycx mori	684	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
612	6	1.2	254	1	LP3_BOMMO	O60506 cricetus	685	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
613	6	1.2	254	1	PRI0_CRIGR	O60468 cricetus	686	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
614	6	1.2	254	1	PRI0_CRIGR	O60468 cricetus	687	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
615	6	1.2	254	1	PRI0_MESAV	P04273 mesocricetu	688	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
616	6	1.2	254	1	PRI0_MOUSE	P04925 mus musculu	689	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
617	6	1.2	254	1	PRI0_SIGHT	P13853 ratus norv	690	6	1.2	266	1	ADP5_MOUSE	O05775 saccharomyc
						O92013 sigmodon hi							

399	6	1.2	151	1	VG10_HSV1	000134	ictaluriid h	472	6	1.2	192	1	RAC2_HUMAN	P15133	homo sapien
400	6	1.2	152	1	CU64_LOCM1	P11739	locusta miq	473	6	1.2	192	1	RAC2_MOUSE	P05144	mus muscunlu
401	6	1.2	154	1	PEPD2_HUMAN	Q9uhv9	homo sapien	474	6	1.2	193	1	HS72_CANAL	P46518	candida alb
402	6	1.2	154	1	PIAS_ORYSA	P20423	oryza sativ	475	6	1.2	194	1	C24A_HUMAN	P13498	h cytochrom
403	6	1.2	154	1	RL2B_FRIAT	O22644	fitilliaris	476	6	1.2	194	1	YG48_AERPE	O9ydb9	aeropyrum p
404	6	1.2	155	1	BCCP_HAETN	P43874	haemophilus	477	6	1.2	195	1	BET1_ECOLI	P17446	escherichia
405	6	1.2	156	1	CU55_ARADI	P80518	araneus dia	478	6	1.2	196	1	TORB_MERTH	O77881	methanobact
406	6	1.2	157	1	CU63_LOCM1	P45584	locusta miq	479	6	1.2	196	1	Y927_TREPA	O83887	treponema p
407	6	1.2	157	1	HMAA_SCHGR	P29556	stictocerc	480	6	1.2	196	1	YEE9_STRCO	O9fbb0	streptomyce
408	6	1.2	158	1	NIFX_AZVOI	P14887	azotobacter	481	6	1.2	198	1	CH19_DPROV	P24563	drosophiila
409	6	1.2	159	1	ECPA_ETKOC	P35645	etkenella c	482	6	1.2	198	1	VIF_BIV06	P19563	bovine immu
410	6	1.2	161	1	IF51_CAEEL	P24563	caenorhabdi	483	6	1.2	199	1	SOD2_HALVO	Q03301	halobacteri
411	6	1.2	161	1	IF52_CAEEL	Q20751	caenorhabdi	484	6	1.2	199	1	SP25_DROME	Q9vvy2	drosophiila
412	6	1.2	161	1	PHAI_HALMA	Q52447	haloarcula	485	6	1.2	199	1	SOD1_HALVO	Q03300	halobacteri
413	6	1.2	162	1	PHAI_PSEMA	O52447	pseudanabae	486	6	1.2	200	1	SDC2_HUMAN	P34900	rattus norv
414	6	1.2	162	1	PHA3_FPRDI	P14876	fremwella d	487	6	1.2	201	1	SDC2_RAT	P34900	rattus norv
415	6	1.2	162	1	RL11_HALNI	P05969	halobacteri	488	6	1.2	201	1	SDC2_MOUSE	P34307	mus muscunlu
416	6	1.2	163	1	CU38_LOCM1	P04375	locusta miq	489	6	1.2	202	1	Y214_PYRHO	O57953	pyrococcus
417	6	1.2	163	1	SODM_MYCKA	P50912	mycobacteri	490	6	1.2	202	1	YM20_PYRAB	Q9v274	pyrococcus
418	6	1.2	163	1	YE99_MYCPN	P75288	mycoplasma	491	6	1.2	202	1	CHPE_PSEAE	O87005	pseudomonas
419	6	1.2	165	1	NG3_DROME	P40140	dtrosophila	492	6	1.2	203	1	PSB3_HUMAN	P49720	homo sapien
420	6	1.2	165	1	TPX_PSEAE	P57668	pseudomonas	493	6	1.2	205	1	PSB3_MOUSE	O9rlp1	mus muscunlu
421	6	1.2	166	1	RNH2_LACIC	O30415	lactococcus	494	6	1.2	205	1	PSB3_ONCMY	O73817	oncothynechu
422	6	1.2	166	1	VP06_BPRED	P27379	bacterioph	495	6	1.2	205	1	PSB3_RAT	P40112	rattus norv
423	6	1.2	167	1	LKC3_PASHA	P55120	pasteurella	496	6	1.2	206	1	C551_CHLYI	P24226	chlorobium
424	6	1.2	167	1	RL9_CHLMU	Q46398	chlamydia m	497	6	1.2	206	1	HS5_MYCLE	Q9x750	mycobacteri
425	6	1.2	167	1	RL9_CHLMR	O84809	chlamydia t	498	6	1.2	206	1	PYRE_BACCL	P46534	bacillus ca
426	6	1.2	167	1	VP6_BPPHE	P11128	bacterioph	499	6	1.2	207	1	SODF_MYCTU	P17670	mycobacteri
427	6	1.2	167	1	YBL7_SCHPO	Q9urva	schizosacch	500	6	1.2	207	1	URK_UREPA	O9pqf9	ureaplasma
428	6	1.2	167	1	YK63_YEAST	P35725	saccharomyc	501	6	1.2	207	1	FCOA_MEREX	O41335	methylobact
429	6	1.2	168	1	BIP3_TOBAC	Q03683	nicotiana t	502	6	1.2	208	1	YJ79_AOUAE	O67788	aquilex aeo
430	6	1.2	168	1	MB12_MYCTU	O05822	mycobacteri	503	6	1.2	208	1	CLPP_CAUCR	O87706	caulobacter
431	6	1.2	168	1	TCTP_ARATH	P31265	arabidopsis	504	6	1.2	209	1	NCRD_VIBCH	O9x496	vibriu chol
432	6	1.2	168	1	TCTP_HORVU	O9m593	hordeum vul	505	6	1.2	210	1	ISPD_CHLPP	O9x7x5	chlamydia p
433	6	1.2	172	1	RL10_AGRIS	O8ue06	agrobacteri	506	6	1.2	211	1	SODP_ORYSA	P93407	oryza sativ
434	6	1.2	172	1	RL10_CHLMU	O9pk78	chlamydia m	507	6	1.2	211	1	T2D9_HUMAN	Q15544	homo sapien
435	6	1.2	172	1	RL10_RHIME	Q92qk9	rhizobium m	508	6	1.2	211	1	RERC_ARATH	O9xw17	arabidopmas
436	6	1.2	172	1	VG16_BPML5	O05222	mycobacteri	509	6	1.2	213	1	CLP1_PSEAE	O912u1	pseudomonas
437	6	1.2	172	1	YFGH_ECOLI	P76572	escherichia	510	6	1.2	213	1	GRPE_DROME	P46604	dtrosophila
438	6	1.2	173	1	APP7_KLUIA	O13350	kluyveromyc	511	6	1.2	213	1	CCMA_RHOCA	P29959	rhodobacter
439	6	1.2	173	1	ATP7_YEAST	P30902	saccharomyc	512	6	1.2	214	1	KAD_ECOLI	P05082	escherichia
440	6	1.2	174	1	CO1A_TENNO	P80681	tenebrio mo	513	6	1.2	214	1	KAD_SALTY	P37407	salmoneila
441	6	1.2	175	1	MP54_PHAHO	P56167	phalaris ag	514	6	1.2	215	1	CASI_RABTY	P09115	oryctolagus
442	6	1.2	175	1	YF49_MYCTU	O10777	mycobacteri	515	6	1.2	215	1	V457_ANASP	P29713	anabaena sp
443	6	1.2	177	1	UCRI_RICPR	O9z495	rickettsia	516	6	1.2	215	1	VEGA_CHICK	P52582	gallus gall
444	6	1.2	179	1	RRE_CHLMU	P71148	chlamydia m	517	6	1.2	216	1	YOFN_BACSU	P54471	bacillus su
445	6	1.2	179	1	RRE_CHLTR	O84684	chlamydia t	518	6	1.2	216	1	COO7_HUMAN	O99807	homo sapien
446	6	1.2	179	1	UCRI_RHOVI	P81380	rhodopsendo	519	6	1.2	217	1	DIIP_HUMAN	Q10837	mycobacteri
447	6	1.2	179	1	YPU1_BACSU	P35156	bacillus su	520	6	1.2	217	1	COT1_MYCTU	O9nyx4	homo sapien
448	6	1.2	180	1	KCV2_BORBU	O51759	borrelia bu	521	6	1.2	217	1	BS25_BOVIN	O27957	homo sapien
449	6	1.2	180	1	Y335_TREPA	O83355	treponema p	522	6	1.2	218	1	ATPL_METJA	O57674	methanococc
450	6	1.2	180	1	YCBU_ECOLI	P75859	escherichia	523	6	1.2	220	1	DEOC_STRPN	O97rh2	streptococc
451	6	1.2	181	1	OLEO_HELAN	P29529	heliianthus	524	6	1.2	220	1	MINC_VIBCH	O9qgn9	vibriu chol
452	6	1.2	182	1	FMIA_ECOLI	P04128	escherichia	525	6	1.2	220	1	NAMD_YERPE	O8sd91	yersinia pe
453	6	1.2	182	1	SRPB_SYNP7	O55026	synchococc	526	6	1.2	220	1	YPY1_ECTHA	P42515	ecetolthorfo
454	6	1.2	183	1	APR_METJA	O59049	methanococc	527	6	1.2	220	1	YC24_GALSU	P435912	galidieria s
455	6	1.2	183	1	RBP4_MYCTU	P71614	mycobacteri	528	6	1.2	221	1	YGB1_HAETN	P44290	haemophilus
456	6	1.2	184	1	ADY_HUMAN	P10109	homo sapien	529	6	1.2	221	1	FLA5_PYRAB	O9uy14	pyrococcus
457	6	1.2	185	1	CASI_HUMAN	P47710	homo sapien	530	6	1.2	222	1	FTSE_ECOLI	P10115	escherichia
458	6	1.2	185	1	PEP3_DROME	O9vyp6	dtrosophila	531	6	1.2	222	1	TPIS_CERYT	O74004	centaethaeum
459	6	1.2	186	1	SODC_PASMU	O59689	pasteurella	532	6	1.2	222	1	TSN2_HUMAN	O60636	homo sapien
460	6	1.2	187	1	KAD_XYIFA	O9pmm3	xyellella fas	533	6	1.2	222	1	YJ73_SYNY3	P73933	synchocyst
461	6	1.2	187	1	RUS2_THIFE	P24330	thiobacillu	534	6	1.2	222	1	SPZR_BACSU	P39157	halobacteri
462	6	1.2	187	1	SODC_HAETN	P25641	haemophilus	535	6	1.2	224	1	VARD_HALNI	O9m125	homo sapien
463	6	1.2	188	1	YUDE_LACIC	P22447	lactococcus	536	6	1.2	225	1	R821_HUMAN	P29967	paracoccus
464	6	1.2	189	1	NISM_BOVIN	Q02380	bos taurus	537	6	1.2	225	1	C553_PARDE	P13003	lactococcus
465	6	1.2	189	1	NISM_HUMAN	O43674	homo sapien	538	6	1.2	226	1	LVCV_BPPHV	O47950	neutrospora
466	6	1.2	190	1	HIN_SALAE	Q02869	salmoneila	539	6	1.2	226	1	NURM_NEUCR	P37884	mesocricetu
467	6	1.2	190	1	HIN_SALTY	P03013	paracoccus	540	6	1.2	226	1	RLI_TREPA	O83366	treponema p
468	6	1.2	190	1	UCRI_PARDE	P05417	paracoccus	541	6	1.2	226	1	T2M6_METJA	P81326	methanococc
469	6	1.2	192	1	PSR3_BOVIN	P33672	bos taurus	542	6	1.2	226	1	RPE_CHLMU	O9pxr7	chlamydia m
470	6	1.2	192	1	RAC2_BOVIN	O9tuz5	bos taurus	543	6	1.2	226	1			
471	6	1.2	192	1	RAC2_CAVPO	O88931	cavia porce	544	6	1.2	228	1			

Tue Jan 28 10:21:48 2003

us-09-889-314-2.rsp

253	7	1.4	1199	1	N121_RAT	P52591	rattus norv	326	6	1.2	113	1	RLA2_DROME	P05389	drosophila
254	7	1.4	1221	1	TOP2_TRYB	P12531	trypanosoma	327	6	1.2	113	1	RLA2_EUPRA	09g02	euplotes ra
255	7	1.4	1224	1	PER_DROME	P07663	drosophila	328	6	1.2	113	1	SU11_MOUSE	P48024	mus musculus
256	7	1.4	1249	1	TPP2_RAT	064560	rattus norv	329	6	1.2	113	1	VC71_MOUSE	Q11048	mycobacteri
257	7	1.4	1276	1	GIL1_ENTRI	P32022	entamoeba h	330	6	1.2	114	1	RBEA_LISMO	09y774	listeria mo
258	7	1.4	1285	1	SL17_ENTRI	P33502	entamoeba h	331	6	1.2	114	1	RK22_CYAPA	P15768	cytomech
259	7	1.4	1324	1	IRS2_HUMAN	06y482	homo sapien	332	6	1.2	114	1	RLA2_BRUMA	09y703	brugia mala
260	7	1.4	1353	1	POLN_DROME	P25822	drosophila	333	6	1.2	114	1	RLA2_ENTTE	P09709	eimeria ten
261	7	1.4	1379	1	Y208_MYCTU	P36286	san miquel	334	6	1.2	114	1	Y467_BOCAL	RLA2-PAR	
262	7	1.4	1901	1	MIS_AEGIR	053553	mycobacteri	335	6	1.2	116	1	INS_LOPEI	INS_LOPEI	
263	7	1.4	1938	1	MC3A_HUMAN	P24733	aquebacter	336	6	1.2	116	1	RLA2_BRAFL	Y318_RICPR	
264	7	1.4	1980	1	MC3A_HUMAN	060318	homo sapien	337	6	1.2	117	1	NIT5_ORYSA	NIT5_ORYSA	
265	7	1.4	2175	1	HMCU_DROME	P10180	drosophila	338	6	1.2	117	1	RL7_CLOPE	RL7_CLOPE	
266	7	1.4	2564	1	SPCO_HUMAN	09h254	homo sapien	339	6	1.2	121	1	RL7_HAEIN	RL7_HAEIN	
267	7	1.4	4196	1	DYHC_SCHPO	013290	schizosacch	340	6	1.2	122	1	YL15_ADE41	YL15_ADE41	
268	7	1.4	4644	1	DYHC_RAT	P38650	rattus norv	341	6	1.2	122	1	RL7_NEIPE	RL7_NEIPE	
269	7	1.2	24	1	GABE_RANRU	P80400	rana rugosa	342	6	1.2	123	1	RL7_RHOSH	RL7_RHOSH	
270	6	1.2	40	1	DLDH_SOLTU	P80403	solanum tub	343	6	1.2	123	1	YC69_MYCTU	YC69_MYCTU	
271	6	1.2	42	1	COAT_BPHLU	007485	bacterioph	344	6	1.2	124	1	TH1H_PICMA	TH1H_PICMA	
272	6	1.2	55	1	RL32_VIRMA	09i336	vlb10 mar	345	6	1.2	124	1	PAND_CAMJE	PAND_CAMJE	
273	6	1.2	55	1	RL32_VIRMA	09i336	vlb10 mar	346	6	1.2	124	1	RL7_STRYG	RL7_STRYG	
274	6	1.2	60	1	NXS1_ASTST	P01438	astrotina st	347	6	1.2	125	1	YH91_AOUAE	YH91_AOUAE	
275	6	1.2	61	1	NXS2_ENHSC	P25493	enhydrina st	348	6	1.2	125	1	RL7_AOUAE	RL7_AOUAE	
276	6	1.2	61	1	PE19_MOUSE	P48539	homo sapien	349	6	1.2	126	1	PHE3_RHOS2	PHE3_RHOS2	
277	6	1.2	62	1	NXSB_LATCR	P07734	mus musculu	350	6	1.2	127	1	RL7_STRAT	RL7_STRAT	
278	6	1.2	62	1	NXSB_LATCR	P25496	laticauda c	351	6	1.2	128	1	YD65_MYCTU	YD65_MYCTU	
279	6	1.2	67	1	RL23_AERPE	P10459	laticauda l	352	6	1.2	128	1	RL7_STRAT	RL7_STRAT	
280	6	1.2	68	1	RL23_AERPE	P27452	streptococc	353	6	1.2	128	1	RL7_STRAT	RL7_STRAT	
281	6	1.2	69	1	RL23_AERPE	Q9y416	aeropyrum p	354	6	1.2	128	1	YD65_MYCTU	YD65_MYCTU	
282	6	1.2	71	1	RL23_AERPE	P32412	archaeoglob	355	6	1.2	129	1	RL7_CHLMU	RL7_CHLMU	
283	6	1.2	71	1	RL23_AERPE	P32412	archaeoglob	356	6	1.2	129	1	RL7_CHLMU	RL7_CHLMU	
284	6	1.2	71	1	RL23_AERPE	P00840	zea mays (m	357	6	1.2	129	1	RL7_CHLMU	RL7_CHLMU	
285	6	1.2	74	1	ATP9_MAIZE	P26855	marichantia	358	6	1.2	131	1	RS12_DEIRA	RS12_DEIRA	
286	6	1.2	74	1	ATP9_MAIZE	P05717	pisum sativ	359	6	1.2	131	1	RS12_DEIRA	RS12_DEIRA	
287	6	1.2	74	1	ATP9_MAIZE	P13547	litticum ae	360	6	1.2	132	1	RS8_CHLTR	RS8_CHLTR	
288	6	1.2	74	1	ATP9_MAIZE	018881	metacac radi	361	6	1.2	133	1	TVB2_HUMAN	TVB2_HUMAN	
289	6	1.2	75	1	TRBK_AGRU	Q9fms3	eschierichia	362	6	1.2	133	1	ANFB_HUMAN	ANFB_HUMAN	
290	6	1.2	75	1	TRBK_AGRU	P54912	agrobacteri	363	6	1.2	133	1	ANFB_HUMAN	ANFB_HUMAN	
291	6	1.2	76	1	APM2_HUMAN	015847	homo sapien	364	6	1.2	134	1	CU3A_TENNO	CU3A_TENNO	
292	6	1.2	76	1	APM2_HUMAN	Q97084	streptococc	365	6	1.2	134	1	VA2_BPT5	VA2_BPT5	
293	6	1.2	79	1	DLTC_STRPN	P20468	rattus norv	366	6	1.2	135	1	RS8_AERPE	RS8_AERPE	
294	6	1.2	85	1	SELW_HUMAN	015532	homo sapien	367	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
295	6	1.2	86	1	SELW_HUMAN	095K14	sus scrofa	368	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
296	6	1.2	86	1	SELW_HUMAN	015532	homo sapien	369	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
297	6	1.2	88	1	SELW_HUMAN	015532	homo sapien	370	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
298	6	1.2	88	1	SELW_HUMAN	015532	homo sapien	371	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
299	6	1.2	88	1	SELW_HUMAN	015532	homo sapien	372	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
300	6	1.2	89	1	SELW_HUMAN	015532	homo sapien	373	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
301	6	1.2	89	1	SELW_HUMAN	015532	homo sapien	374	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
302	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	375	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
303	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	376	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
304	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	377	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
305	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	378	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
306	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	379	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
307	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	380	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
308	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	381	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
309	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	382	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
310	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	383	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
311	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	384	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
312	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	385	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
313	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	386	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
314	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	387	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
315	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	388	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
316	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	389	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
317	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	390	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
318	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	391	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
319	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	392	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
320	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	393	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
321	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	394	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
322	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	395	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
323	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	396	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
324	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	397	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	
325	6	1.2	91	1	SELW_HUMAN	015532	homo sapien	398	6	1.2	136	1	HNS_ECOLI	HNS_ECOLI	

107	1.4	300	1	FR12_MAIZE	P29390	zea mays (m	180	7	1.4	482	1	BTBL_HUMAN	Q9H0C5	homo sapien
108	1.4	304	1	MTRE_METMA	P80651	methanosarc	181	7	1.4	482	1	MKRL_HUMAN	Q9UNC7	homo sapien
109	1.4	306	1	CH38_DROME	P07183	drosophila	182	7	1.4	483	1	KPYR_METEX	Q05118	methylobact
110	1.4	309	1	YEEY_ECOLI	P76369	eschlerichia	183	7	1.4	486	1	CDSN_HUMAN	Q15517	homo sapien
111	1.4	311	1	TRXB_MTCM	O30973	mycobacteri	184	7	1.4	488	1	BTBL_MOUSE	P58544	mus musculu
112	1.4	313	1	BPB7_ECOLI	P33364	eschlerichia	185	7	1.4	488	1	NOM2_PSEAE	Q9HTR0	pseudomonas
113	1.4	319	1	K6PF_CLOPE	O8XNH2	clostridium	186	7	1.4	504	1	A37C_DROME	O18477	drosophila
114	1.4	321	1	TRXB_STRO	O05741	streptomyces	187	7	1.4	530	1	ZIC2_MOUSE	O62820	mus musculu
115	1.4	321	1	TRXB_STRO	P52215	streptomyces	188	7	1.4	535	1	VHTJ_BPT3	P20323	bacterioph
116	1.4	322	1	ADT_SCHPO	O09188	schizosacch	189	7	1.4	536	1	YEN1_SCHPO	P20323	bacterioph
117	1.4	323	1	JUND_CHICK	P27921	gallus gall	190	7	1.4	541	1	YA97_MYCPN	O13365	schizosacch
118	1.4	324	1	G3P1_GIRO	O16027	globoidea r	191	7	1.4	545	1	H15_DROME	Q94690	drosophila
119	1.4	328	1	HXDL_HUMAN	O94220	homo sapien	192	7	1.4	545	1	ARAB_STAAM	Q99457	staphylococ
120	1.4	331	1	FLIG_ECOLI	P31067	eschlerichia	193	7	1.4	548	1	HEM1_YEAST	P09950	saccharomyc
121	1.4	331	1	FLIG_SALT	P15933	salmonella	194	7	1.4	556	1	SYR_LISIN	Q92712	listeria in
122	1.4	333	1	TRBL_ARATH	O39243	arabidopsis	195	7	1.4	556	1	SYR_LISMO	O84493	listeria mo
123	1.4	339	1	HXDD_MOUSE	P70217	mus musculu	196	7	1.4	565	1	CNIA_MOUSE	O61481	mus musculu
124	1.4	344	1	YD33_MYCTU	O10644	mycobacteri	197	7	1.4	566	1	CRT1_ARATH	Q07356	arabidopsis
125	1.4	346	1	ASPG_HUMAN	P20933	homo sapien	198	7	1.4	566	1	CRT1_ORYSA	O94006	oryza sativ
126	1.4	346	1	ASPG_MOUSE	O64191	mus musculu	199	7	1.4	570	1	CRT1_NARPS	O40406	narcissus p
127	1.4	351	1	MURG_HAEIN	P45065	haemophilus	200	7	1.4	571	1	CRT1_SOYBN	P28553	glycine max
128	1.4	354	1	MURG_PASMO	P57817	pasteurella	201	7	1.4	571	1	CRT1_MAIZE	P49086	zea mays (m
129	1.4	358	1	ALF_EDWIC	O52402	edwardsiella	202	7	1.4	588	1	CMC2_CAEEL	Q20799	caenorhabdl
130	1.4	359	1	CKR3_MOUSE	P51678	mus musculu	203	7	1.4	606	1	PF20_CHLRE	P93107	chlamydomon
131	1.4	364	1	YPH2_THIIV	P45367	thiocystis	204	7	1.4	610	1	DNAK_METTE	Q9UXR0	methanosarc
132	1.4	365	1	CAD1_ARATH	P48523	arabidopsis	205	7	1.4	627	1	DNAK_GUTH	P29315	guillardia
133	1.4	370	1	PCP1_AMPCA	P80484	amphidipnis	206	7	1.4	632	1	PAPB_DROME	P21187	drosophila
134	1.4	375	1	Y112_UREPA	O9P332	ureaplasma	207	7	1.4	637	1	ODP2_AZOVI	P10802	azotobacter
135	1.4	376	1	NIR_AUCEA	P38501	alcaligenes	208	7	1.4	637	1	YHE7_YEAST	P38724	saccharomyc
136	1.4	378	1	NIR_ACHCY	P25006	achromobact	209	7	1.4	648	1	RAD4_SCHPO	P32372	schizosacch
137	1.4	379	1	NIR_PSESP	O01537	pseudomonas	210	7	1.4	649	1	GDMD_YEAST	P32191	saccharomyc
138	1.4	382	1	NTRB_RHIME	O52977	rhizobium m	211	7	1.4	654	1	NUOL_STROCO	O9XAR5	streptomyces
139	1.4	383	1	TRB2_ARATH	O39242	arabidopsis	212	7	1.4	655	1	PER_DROST	O03355	drosophila
140	1.4	384	1	GAS1_MOUSE	O01721	mus musculu	213	7	1.4	657	1	PRIA_TREPA	O83258	treponema p
141	1.4	388	1	TRB1_PPRAB	O9V198	pyrococcus	214	7	1.4	657	1	Y023_NPVOP	O10282	orygia pseu
142	1.4	389	1	TRPB_PYROK	O9YB0	pyrococcus	215	7	1.4	661	1	PER_DROSE	O03354	drosophila
143	1.4	390	1	FLAA_BORBR	O06064	bordeletia	216	7	1.4	676	1	PER_DROMA	O03354	drosophila
144	1.4	391	1	SOX1_MOUSE	P53783	mus musculu	217	7	1.4	708	1	KHS1_YEAST	P36690	saccharomyc
145	1.4	392	1	PORA_THEMEA	O05651	thermotoga	218	7	1.4	728	1	YMX6_CAEEL	P34514	caenorhabdl
146	1.4	396	1	PORA_PYRPU	O51804	pyrococcus	219	7	1.4	733	1	HIC1_HUMAN	O14526	homo sapien
147	1.4	398	1	CSP_PLACG	P08673	plasmodium	220	7	1.4	749	1	PRIA_CHLUP	O926Y2	chlamydia p
148	1.4	401	1	CSP_PLACG	P08674	plasmodium	221	7	1.4	753	1	PRIA_CHLMO	O926Y2	chlamydia m
149	1.4	404	1	YF82_MYCLE	O9C94	mycobacteri	222	7	1.4	753	1	PRIA_CHLTR	O484783	chlamydia t
150	1.4	409	1	TRPB_METVO	P14638	methanococ	223	7	1.4	778	1	YFK8_YEAST	P43410	saccharomyc
151	1.4	411	1	YGA1_YEAST	P42942	saccharomyc	224	7	1.4	778	1	YQ34_MYCTU	P71933	mycobacteri
152	1.4	413	1	ACDS_PIG	P79273	scrofa	225	7	1.4	779	1	K6PF_HUMAN	P08837	homo sapien
153	1.4	416	1	EURL_CAEEL	P28515	caenorhabdl	226	7	1.4	779	1	K6PF_MOUSE	P47857	mus musculu
154	1.4	419	1	COOB_ECOLI	P25325	eschlerichia	227	7	1.4	779	1	K6PF_RABIT	P00511	oryctolagus
155	1.4	420	1	ZG58_XENLA	O8VHW5	rattus norv	228	7	1.4	779	1	K6PF_RAT	P47858	rattus norv
156	1.4	421	1	CCG8_RAT	O8VHW5	rattus norv	229	7	1.4	781	1	K6PF_CANFA	P52784	cantis fami
157	1.4	423	1	CCG8_MOUSE	O8VHW5	mus musculu	230	7	1.4	803	1	GYRB_BUCAT	P57126	buchnera ap
158	1.4	425	1	LBH2_HUMAN	P08810	homo sapien	231	7	1.4	803	1	GYRB_BUCAP	P29435	buchnera ap
159	1.4	426	1	VG4_BPPI	P03666	bacterioph	232	7	1.4	805	1	PRIA_BACSU	P94461	bacillus su
160	1.4	426	1	VG4_BPPI	P03666	bacterioph	233	7	1.4	826	1	RSG5_HUMAN	O43374	homo sapien
161	1.4	435	1	EF1A_SULTO	O976B1	sulfolobus	234	7	1.4	831	1	SORT_HUMAN	O99523	homo sapien
162	1.4	436	1	BT1_MAIZE	P29518	zea mays (m	235	7	1.4	852	1	WS14_HUMAN	O99523	homo sapien
163	1.4	445	1	HH3R_CAVPO	O9J135	cavia porce	236	7	1.4	856	1	CLPB_HELPJ	O9ZBH1	helicobacte
164	1.4	445	1	HH3R_CAVPO	O9J135	cavia porce	237	7	1.4	856	1	CLPB_HELPJ	P71404	helicobacte
165	1.4	445	1	HH3R_HUMAN	P58406	mus musculu	238	7	1.4	866	1	ODP1_HAEIN	P45119	haemophilus
166	1.4	445	1	HH3R_MOUSE	O9QYN8	rattus norv	239	7	1.4	892	1	HIC1_MOUSE	O914Y5	mus musculu
167	1.4	452	1	NORM_BACSU	O31855	dictyostell	240	7	1.4	892	1	TNPA_STYFR	P20189	streptomyces
168	1.4	457	1	TBA_DICDI	P32255	dictyostell	241	7	1.4	923	1	YAUA_SCHPO	O10165	schizosacch
169	1.4	457	1	UIDB_ECOLI	P30868	eschlerichia	242	7	1.4	924	1	ICAS_HUMAN	O9UNF0	homo sapien
170	1.4	458	1	ODP2_NEUCR	P20285	neutrospora	243	7	1.4	931	1	TRPE_HUMAN	O9Y210	homo sapien
171	1.4	458	1	TRXB_MYCLE	P46643	mycobacteri	244	7	1.4	941	1	MSH1_SCHPO	O13321	schizosacch
172	1.4	465	1	SNX8_HUMAN	O9YX2	homo sapien	245	7	1.4	942	1	M3KE_MOUSE	O9W161	mus musculu
173	1.4	470	1	YEE7_YEAST	P32614	saccharomyc	246	7	1.4	988	1	OMB_DROME	O24432	drosophila
174	1.4	473	1	NORB_PSEST	P98008	pseudomonas	247	7	1.4	992	1	POLS_RUBVM	O08653	rubella vir
175	1.4	474	1	CRT1_SYNP7	P26294	synechococ	248	7	1.4	1029	1	YK95_RHIME	O52899	rhicobium m
176	1.4	474	1	FLA_LRCMI	P34606	legionella	249	7	1.4	1063	1	POLS_RUBVM	P23180	rubella vir
177	1.4	476	1	YAAJ_ECOLI	P30143	eschlerichia	250	7	1.4	1078	1	CARB_BUCAT	P57244	buchnera ap
178	1.4	478	1	VP26_DROME	Q9W552	drosophila	251	7	1.4	1171	1	DPOI_HSVT2	Q9YU52	herpesvirs
179	1.4	482	1	BHB3_HUMAN	Q9C0J9	homo sapien	252	7	1.4	1189	1	POL_BAEVM	P10272	baboon endo

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:33:12 : Search time 20 Seconds  
(without alignments)  
1028.613 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496

Sequence: 1 DTMMSSSSSGPNOKNIMS.....LKAYALSGALGAHKTNMF 496

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_40:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.2	126	1	RL7_RHIME
2	11	2.2	674	1	PRGA_CORGL
3	11	2.2	2038	1	FSH_DROME
4	9	1.8	88	1	RS21_RHITO
5	9	1.8	111	1	RLA4_CLAHE
6	9	1.8	169	1	PKBS_HUMAN
7	9	1.8	507	1	CRT1_STRGR
8	9	1.8	508	1	CRT1_STRSE
9	9	1.8	747	1	ELS_BOVIN
10	9	1.8	782	1	PRTP_HSVT2
11	9	1.8	860	1	ELS_MOUSE
12	9	1.8	967	1	SOXA_CORSL
13	9	1.8	1001	1	ORK1_DROME
14	9	1.6	109	1	VATL_MAIZE
15	8	1.6	110	1	RLA2_SCCHO
16	8	1.6	110	1	RLA2_SCCHO
17	8	1.6	111	1	RLA2_ARTSA
18	8	1.6	111	1	RLA3_CLAHE
19	8	1.6	112	1	RLA3_TRYCR
20	8	1.6	132	1	Y615_AOUAE
21	8	1.6	144	1	NR01_CORAM
22	8	1.6	161	1	VATL_NEUCR
23	8	1.6	161	1	VATL_SCCHO
24	8	1.6	165	1	VATL_AVEBSA
25	8	1.6	165	1	VATL_ORYSA
26	8	1.6	205	1	GSCL_HUMAN
27	8	1.6	215	1	RL10_EUGER
28	8	1.6	272	1	VG67_HSEVB
29	8	1.6	311	1	RPAS_ECOLI
30	8	1.6	338	1	PSBO_EUGER
31	8	1.6	344	1	LIC2_BURCE
32	8	1.6	361	1	HYPB_MYCEU
33	8	1.6	367	1	TRMU_MYCTU

34	8	1.6	460	1	MOC3_HUMAN
35	8	1.6	472	1	CRT1_SYNR3
36	8	1.6	479	1	ZDS_ANASP
37	8	1.6	512	1	LEU1_LISTIN
38	8	1.6	512	1	LEU1_LISTMO
39	8	1.6	531	1	V4T0_RHISN
40	8	1.6	594	1	COX1_PHYPO
41	8	1.6	631	1	Z179_RAT
42	8	1.6	632	1	Z179_HUMAN
43	8	1.6	687	1	SPD1_NEPCL
44	8	1.6	747	1	YN54_CAEBL
45	8	1.6	747	1	ATX7_HUMAN
46	8	1.6	892	1	YBE7_YEAST
47	8	1.6	1381	1	YBAS_ECOLI
48	8	1.6	1534	1	41_DROME
49	8	1.6	1698	1	RS16_HELPY
50	7	1.4	76	1	YPJK_ECOLI
51	7	1.4	78	1	RS18_CAUCR
52	7	1.4	92	1	SASG_BACCE
53	7	1.4	93	1	RLA3_YEAST
54	7	1.4	106	1	RLA1_ARATH
55	7	1.4	110	1	RLA1_CLAHE
56	7	1.4	110	1	RLA4_YEAST
57	7	1.4	111	1	RLA2_BABBO
58	7	1.4	111	1	RLA2_ALTAL
59	7	1.4	113	1	DAD1_PRA
60	7	1.4	117	1	RL7_LISMO
61	7	1.4	120	1	RL7_STRPY
62	7	1.4	121	1	RL7_BACSU
63	7	1.4	122	1	RL7_HALPR
64	7	1.4	122	1	RL7_CIOAB
65	7	1.4	123	1	RL7_BRIME
66	7	1.4	124	1	RL7_RHITO
67	7	1.4	125	1	RL7_RITON
68	7	1.4	125	1	RL7_CAUCR
69	7	1.4	127	1	CH16_DROGR
70	7	1.4	142	1	HBH_LATCH
71	7	1.4	146	1	SSI_STRYL
72	7	1.4	146	1	YEVO_YEAST
73	7	1.4	148	1	FM12_PSEAE
74	7	1.4	154	1	FMK1_PSEAE
75	7	1.4	154	1	PLAS_STLPR
76	7	1.4	165	1	NPM3_MOUSE
77	7	1.4	175	1	DNLE_RAPSA
78	7	1.4	184	1	RNFG_SALTI
79	7	1.4	206	1	2137_HUMAN
80	7	1.4	207	1	NORD_HARIN
81	7	1.4	208	1	PCP_LACIC
82	7	1.4	215	1	VG32_BPMU5
83	7	1.4	220	1	HEX8_ADE02
84	7	1.4	225	1	HEX8_ADE05
85	7	1.4	227	1	VG07_BPPT2
86	7	1.4	229	1	PYRF_STAMM
87	7	1.4	230	1	HEX8_ADE12
88	7	1.4	233	1	Y057_CAEBL
89	7	1.4	235	1	PLIG_SHIBO
90	7	1.4	242	1	TRT2_ECOLI
91	7	1.4	243	1	TRT4_ECOLI
92	7	1.4	244	1	TRT1_ECOLI
93	7	1.4	244	1	EGRE_HUMAN
94	7	1.4	247	1	POOC_KLEEN
95	7	1.4	251	1	TPIS_LISTIN
96	7	1.4	251	1	VG38_BPT2
97	7	1.4	252	1	YFC7_CAEEL
98	7	1.4	266	1	DHP2_ECOLI
99	7	1.4	269	1	NIFH_CIOAB
100	7	1.4	271	1	YCDN_ECOLI
101	7	1.4	272	1	KPRA_DEIRA
102	7	1.4	276	1	
103	7	1.4	283	1	
104	7	1.4			
105	7	1.4			
106	7	1.4			

095396 homo sapien  
P29273 synechocyst  
09664 anabena sp  
Q2228 listeria in  
08y59 listeria mo  
P5569 rhizobium s  
Q07434 physarum po  
070418 rattus norv  
09u15 homo sapien  
P51802 rattus norv  
P19837 nephila cla  
P34588 caenorhabdi  
O15265 homo sapien  
P34216 saccharomyc  
P76464 escherichia  
09y89 drosophila  
P56023 helicobacte  
P5213 escherichia  
Q9a7q2 caulobacter  
P07787 bacillus ce  
P10622 saccharomyc  
O23095 arabidopsis  
P50344 cladosporiu  
P02400 saccharomyc  
Q9uuz6 aspergillus  
P27055 babesia bov  
P42037 alternaria  
Q9zra3 plium sativ  
O8yaa3 listeria mo  
P58076 streptococc  
P02394 bacillus su  
P14134 halonacero  
Q97eg8 clostridium  
P41106 bruceella me  
Q98a67 rhizobium l  
Q92187 rickettsia  
Q9aau7 caulobacter  
P24510 drosophila  
P23741 latimeria c  
Q06871 streptomyc  
P40092 saccharomyc  
P17836 pseudomonas  
P07030 silene pret  
Q9c9d0 mus musculu  
P21288 raphanus sa  
Q826q7 salmonella  
P52743 homo sapien  
P43958 haemophilus  
O87765 lactococcus  
Q05241 mycobacteri  
sus scrofa  
P03280 human adeno  
P24936 human adeno  
Q01074 bacteriopho  
Q99u43 staphylococ  
P36713 human adeno  
Q09463 caenorhabdi  
P35745 siligella bo  
P32885 escherichia  
P15177 escherichia  
P13979 escherichia  
Q92915 homo sapien  
P27505 klebsiella  
Q92811 listeria in  
Q8y413 listeria mo  
P07875 bacteriopho  
P06499 simian aden  
Q45731 caenorhabdi  
P19539 escherichia  
Q97me5 clostridium  
P75901 escherichia  
Q9rrr1 deinococcus







```

Db 502 AACAAAGCGTGAACCTGTGTGCGACCCCTGGGTCGCCGAAGAGCGGTGTCACCTGCTC 561
QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyValSerGlnThrLys 42
Db 562 TCTGGGAAGTCAACCAAG-----AAGTCACGACGAG 591
QY 43 GlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAla 62
Db 592 CCCTTGGCAAAACACTGTCTGGGCTCAGAAACTGAGGAGGAGGCAATGCCAAGCCCTC 651
QY 63 GlyAlaSerGlyLys----- 67
Db 652 GGACCCACTGCCAAGCTCGGAAGGTCGACGGCGCCCAAGCAGCAGTCCAGTGAAGAT 711
QY 68 -----AspLysThrSerThrThrThrLysThrGluThrAlaProGln 82
Db 712 TCCTCCATCTCAAGCGATGACAGATGTCGAGTGGAATTCACGAAACACGACCCG 771
QY 83 GlyAlaAlaAlaGlyLysGluSerSerGlu-----SerGln 94
Db 772 GCCAAGGTCGACAGCCCTGCGCAAGATCTCCAGCAAGAACAGCCCGCCCTACC 831
QY 95 LysAlaGly-----AlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsn 112
Db 832 AAGTTAGGGGAATGTGGCGCCACCTGCTTAACACAGCGGCGGACGCGACGCTGCT 891
QY 113 ThrAlaThrLysIleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGluSer 132
Db 892 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
QY 133 ThrLeuGluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAla----- 150
Db 952 GACTCAGACAGTGAAGAGACAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011
QY 151 -----ValValValAlaAlaLeuSerGlyLys----- 159
Db 1012 AAAGTCCCATGTGACAGCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
QY 160 -----SerSerGlySerAlaLysLeuGlnThrPro 169
Db 1072 ATCTTAGACACCCCGAAGAGACTGGGCTGCGACGCCACCCAGCAGGCAAGCAAGGCCA 1131
QY 170 GluLeuProLysProGlyValThrProArgSerGluValIleGluIleGluLeuAlaLeu 189
Db 1132 GAG-----AAGGACTCGGAAACCAAGCAGTGAAGCATTTGATGATGAGATGAAAG 1185
QY 190 AlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsnTyraLaser 209
Db 1186 CCAGTCACTGTGAATATCTCTCAGCAAGAGACTTCTGGGAAGGCCCTCGGCCAGAGGT 1245
QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGluLysGlnAlaIleLysIle 229
Db 1246 ACCCTGACCCCGCCGCAAGAGTATCTCCAGAAAGG-----GCTCTCTCAGTCAACCTT 1299
QY 230 AspLysGluArgGluGluGlnLysAlaAlaGlnLysAlaAlaGlnLysSerLysAspLeu 249
Db 1300 GGAAGGCAAGAGCCGTGTGTCAGCCAGCAGGAAACCAAGCAAGCAAGCAGTGAAG--- 1356
QY 250 GluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAla----- 266
Db 1357 GAGTCAGAGAGTGAAGTGGGAGACACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
QY 267 -----IleThrValIleSerIleValAlaAlaIlePheThrLysGlyAlaGly 282
Db 1417 AAGGTGAACCTTTGGGGAAGAGCTCCAGGTCAAGCTGTTCCACCGTCAACCCCGGGG 1476
QY 283 LeuAlaGlyLeuAlaAlaGly-----AlaAlaValAlaAlaAlaAlaAlaAlaAlaAla 297
Db 1477 TCATCGGGAAGAGGTGCCAAGCTGCGCTGCTGCGGAAGGTGGGTCAAGCTCTCAGG 1536
QY 297 ----- 297
Db 1537 GTCCAAATGTTAAGAAAGAGATGTCTCGAGAGCAGCAGTGCAGAGCTGCAGACTGCAG 1596

```

```

QY 298 GlyGlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAla 317
Db 1597 GGGCTGGAGCCCGACCCAGGCAAGGCTCTCTTGGCTCTCCCTCAGAGGTGAGGCCCT 1656
QY 318 ValValGlnAlaValLys-----GlnAlaVal 326
Db 1657 GTGGCCACCCAGCTCAGACCTCAGACGGGGCAAGGCCACTCAGGAGGAGTGAAGATCA 1716
QY 327 IleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIle 346
Db 1717 TCTGACAGGAGGAGGAGGACACACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
QY 347 LysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSer 366
Db 1777 GAAAGCAGATGAAGCTTCTCTGTAAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1836
QY 367 LysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysVal 386
Db 1837 -----ACCTGCTCCATTTGAAGCAGGCGGTG 1866
QY 387 IleSerSerLeuThrSerLysTrpValThrValGlyValGlyValValAlaAlaPro 406
Db 1867 ACCCTCTCAGCCAGC-----CTGTCATCCCA 1893
QY 407 AlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsn-ValAlaGly 426
Db 1894 GCTTGGCCCAAGGCG-----ACCCAGAGCTGATGTGACTCTTCCAGTGAAGTCT 1944
QY 426 nPheGlnLysGluVal-----GlyLysLeuGlnAlaAla 438
Db 1945 GAGTCAGAAGGAGCTGCTCCAGCAACCCAGGATACAGGAGAGCTGCGGGCAAGGCG 2004
QY 438 aAspMetIleSerMet---PheThrGlnPheTrpGlnGlnAla----- 451
Db 2005 CTCCAAGGGAAGCTGCTTGGGCAAGGCGTGGCCCAAGTGCACACTCAGAAGCAGG 2064
QY 452 -----SerLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMetThrGlnLysAl 469
Db 2065 CCTTCGGTCAAGGCTATGCTCAGGAAGACTCAGAGAGCTCGAGAGGACTCCAGCAGC 2124
QY 469 aThrLys 471
Db 2125 GAGGAAG 2131

```

RESULT 44

AZ312601

LOCUS 526 bp DNA linear GSS 29-SEP-2000

DEFINITION 1M0028104R mouse 10kb plasmid U06C1M library Mus musculus genomic clone U06C1M0028104 R, DNA sequence.

ACCESSION AZ312601

VERSION AZ312601.1 GI:10356717

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 526)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C., Islam,M., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weis,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu



```

Oy 142 AlalaaglnmetlsglvalaigluvalaValaAlaAlaAlaSerGlyLysSer 161
Db 563 AAGAGCAAGCTAAGGAGCAACAGAGATGCTTCAGAGAGGCAAGGCTGACACT 622
Oy 162 GlySerAlaLysLeuGluThrProGluLeuProLysPro----- 174
Db 623 GGGGAGGCGCAAGGTGAAGAGACAGCTGTAGCCCGGAGCAGAGATGCACTGTGCAG 682
Oy 175 -----GlyValaThrProArgSerGluValIleGluIleGlyLeuAlaAlaLys 191
Db 683 GCAGCGATGAGCGCCATGCTATGCGACCATGTGAGAGGTGCAACAGCTGCAGGGCAAG 742
Oy 192 AlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSer----- 205
Db 743 ---ATCCGACTCTTCAGAGAACAGCTGAGAACGCTTAACACCAGCTGGCCGCTTA 799
Oy 206 -----AsnTyrAlaSerThrGlnAlaAla 214
Db 800 CACGAGAGAACTCCATCTTGAAGGATGCTTGAACAGCTACAGCAGCAGCTGAGAAAGC 859
Oy 215 AspGlnThrAsnLysLeuGly---LeuGluLysGlnAlaIleLysIleAspLysGlu--- 232
Db 860 AAGCAAGACAGACAGACTGGCAAGCTCGACAGAGGCTCAGCAACAGTCAACAGAGAGCTA 919
Oy 233 -----ArgGluGluTyrGlnGluGluMetLysAlaAlaGluGlnLys 245
Db 920 GTGGAAGAGTCAAGAGCGCTCCCGGAGAGAGAGCAACAGGAAAGCCCTGGAACCAAG 979
Oy 246 SerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaValSerVal 265
Db 980 GCAGCGACCTTGAAGAAACAGAGCTCT-ACAGTT----- 1011
Oy 266 AlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGly 285
Db 1012 -----GCGAGC 1017
Oy 286 Leu-----AlaAlaGlyAlaAlaValaGlyAlaAlaAlaGly--- 298
Db 1018 ATCCCAACAGAGAGTGAAGAGGCTCTGCAAGAGCGCTTGAGAGAGTCAACCCGAGACT 1077
Oy 299 -----GlyAlaAlaGly 302
Db 1078 CTGCCGGGCAAAAGCAGCCATGCCATCCGAGCGGATGCTGAAAAGGCTCAGAGACA 1137
Oy 303 AlaAlaAlaAlaThrValAlaAlaThrGlnIlePheVal-GlnAlaValaGlnAlaVala 322
Db 1138 GCAGCAGCGAGTGGCAGAGTTCGACAGCAAAATACAGTCATCTGAGGTGAAGTAAAG 1197
Oy 322 LysGlnAlaValIleThrAlaValaLargLAlaIleThrAlaAlaIleLysAlaAlaVala 342
Db 1198 CAAGTGTCAACAGCTAGCTACTGCTCATCGGAGCTCAAGAGAGGCC-----AGGCGGA 1251
Oy 342 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSe 362
Db 1252 GAACCTCAACAGCTCAGAGAGAGATCCGTCATCAACAGCCCTTGTGGAAGAGCGCAGGC 1311
Oy 362 LysGlyIleSerLysValaPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProly 382
Db 1312 CCAGAGACCCAGAGCCAGCCAGCTGAGGCCAACCCAG----- 1348
Oy 382 LysSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValaGlyVala 402
Db 1348 ----- 1348
Oy 402 ValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnG 422
Db 1349 -----CACGAGACTCGCTCAAGAGAGCTGAGACTC 1377
Oy 422 AsnValAlaGlnPheGlnLysGluValaGlyLysLeuGlnAlaAlaAlaAspMetIleSe 442
Db 1378 CCAGGTGCTGCTGCGAAGGAACACAGTAGCTCAAGAGGCTATAGGA----- 1429
Oy 442 mMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGluSe 462

```

```

Db 1430 -----CAACAGAGGGGAAGAA 1446
Oy 462 rAsnGluMetThrGlnLysAlaThrLys 471
Db 1447 CAATGACCTCCGGGAAAGAACTGGAAG 1474
RESULT 42
BC030898 3441 bp mRNA linear HTC 13-JUN-2002
LOCUS Mus musculus, similar to host cell factor C1, clone IMAGE:4502186,
DEFINITION mRNA.
ACCESSION BC030898 GI:21410147
VERSION BC030898.1
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3441)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NHI-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 44 Row: e Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: incomplete processing.
FEATURES
source
1..3441
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4502186"
/rissue_type="Eye, retina, mouse strain C57BL/6"
/clone_id="NHI_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 849 a 969 c 813 g 810 t
ORIGIN
Alignment Scores:
Pred. No.: 0.183 Length: 3441
Score: 126.00 Matches: 73
Percent Similarity: 32.83% Conservative: 58
Best Local Similarity: 18.30% Mismatches: 122
Query Match: 5.32% Indels: 146
DB: 11 Gaps: 10
US-09-889-314-2 (1-496) x BC030898 (1-3441)
Oy 65 SerGlyLysAspLysThrSerSerThrLysThrGluThrAlaProGlnGlnGly--- 83
Db 2 ACAGGTACCCAGCACACAGCACAGGTCACCTTAACATGAGCTCAACCAACCAACCTT 61
Oy 84 ---ValAlaAlaGlyLysGluSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyVal 102

```







Library"this sequence is part of a project of ESM assemblies resulting from the application of public contests to seed Dupont contests; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maizee Mapping Project"



```

misc_feature
    /db_xref="MGD:MG1:1903394"
    /db_xref="Ensembl:10090"
    /clone="4921511116"
    /sex="male"
    /tissue_type="testis"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="adult"
    /note="data source:SPRR, source key:Q9NTH0, evidence:ISS
    homolog to PUTATIVE SERINE-RICH PROTEIN (FRAGMENT)"
    /db_xref="MGD:MG1:1918175"

BASE COUNT   475 a   346 c   391 g   378 t

ORIGIN

Alignment Scores:
Pred. No.:      0.0546      Length:      1590
Score:          126.50      Matches:      93
Percent Similarity: 39.78%      Conservative: 90
Best Local Similarity: 20.22%      Mismatches: 195
Query Match:      5.34%      Indels:      84
DB:               11      Gaps:      17

US-09-889-314-2 (1-496) x AK014874 (1-1590)

QY      23  LeuThrSerThrProGlnGlyValProGlnGlnAspLys-----LeuSerGlyAsn 39
      ::::|||||:  ::::|||||:  ::::|||||:
DB      74  ATATCAATGTCAGTCAATGAGGGTGAAGAGCCATCCAAAAGAGCCGCTAGAGAGAAA 133
QY      40  GluThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAla 59
      ::::|||||:  ::::|||||:  ::::|||||:
DB     134  AACAAATCTTCACGTGAGGAGATCATGGGAAAAATCTGCCAAAAGGATGCTTCAGCA 193
QY      60  ThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAla 79
      ::::|||||:  ::::|||||:  ::::|||||:
DB     194  GCTCAGCAGAAAACAGTTCGCGCTCAGGAGGGGCTCCTCC--ACCAAAATCCGAGAGTGTG 250
QY      80  ProGlnGlnGlyAlaAlaGlyLysGluSerSerGlnSer---GlnLysAlaGlyAla 98
      ::::|||||:  ::::|||||:  ::::|||||:
DB     251  -----GGGACCTGAGTCTGGAGCAGCTCCAGTGGCTCCATCAGGATAGAGCCACA 301
QY      99  AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerSerThrAlaThrLysIleAla 118
      ::::|||||:  ::::|||||:  ::::|||||:
DB     302  AGTGAAGGAGACAGATCTGTGTGACAGCAGAGCAGCAATCTCTCTCAGTAAACCGCA 361
QY     119  -----MetGlnThrSerIleGlnGluAla 126
      ::::|||||:  ::::|||||:  ::::|||||:
DB     362  GGGTCTGAAAAGGCTTGGAGTCAAGATCAACATCAATCAAGTGTTCAGGCTCATTTGTT 421
QY     127  SerLysSerMetGluSerThrLysGlnSerLysSerLysSerLysAlaAlaGlnMetLys 146
      ::::|||||:  ::::|||||:  ::::|||||:
DB     422  TCTTCCCTGTTGAAGTCCATATGATATGATACACCAATCCACTGTAACAGACAA 481
QY     147  GluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerSerSerAlaLysLeu 166
      ::::|||||:  ::::|||||:  ::::|||||:
DB     482  CAGAGGGGCTCACTAAAGGGGCTCTGGAAGGT---TCTTCAAGCTCAGCCTCTCAG 538
QY     167  GluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGlnIleGly 186
      ::::|||||:  ::::|||||:  ::::|||||:
DB     539  AGCAGCTCAAGATTTGAGTGGCTTGTGGTCTCT---CTGAAAGTGCA-GAAGTAGAG 594
QY     187  LeuAlaLeuAlaLysAlaIleGlnThrLysGlnGlyGluAlaThrLysSerAlaLeuSerAsn 206
      ::::|||||:  ::::|||||:  ::::|||||:
DB     595  TTGGCATTTACTG---TCTTGAATCTAGTTTCAGAGACAGCTTCCAGTGGGTTTAACTCT 651
QY     207  TyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGluLysGlnAla 226
      ::::|||||:  ::::|||||:  ::::|||||:
DB     652  AAATCTAGTTCAGAGCAAAACATTTCAATCAATCA----- 684
QY     227  IleLysIleAspLysGluArgGluGluThrGlnGlnMetLysAlaAlaGlnGlnLysSer 246
      ::::|||||:  ::::|||||:  ::::|||||:
DB     685  -----GTTTCAAAAACAGTTCATCTCAGGAGCTCTTACTTAACCCCCCAGAGAGT 738
QY     247  LysAspLeuGlnGluThrMetAspThrValAsnThrValMetIleAlaValSerValAla 266

```

```

DB     739  TCCACAATTCATGCTGCTGATATCCAAAGCAGGCTCAGTAGTGTCTCACTGCTA 798
QY     267  IleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu 286
DB     799  GCGACC----- 804
QY     287  AlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAla 306
DB     805  AAAAGTGGCGCTTCTTGAAGAGCTGTCTCAGTACGTCCAGAGAGTGTCTCAGAGT 864
QY     307  ThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVal 326
DB     865  AGCACT-----TCACAGTGGCTTCGAAAGTACTCCAGGCGCAGAGAGTGTCTGTC 918
QY     327  IleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIle 346
DB     919  AAGTTTGCTTGTGCGAAGTTACCAATGAAACATATAAACAAGAAC----- 963
QY     347  LysAlaPheIleLysThrLeuValLysAlaIleAla---LysAlaIleSerLysGlyIle 365
DB     964  CAACCTTCTTCATATGACTGTATAGACAGGTGGCTGGAAGTTAGTACACTTGTCT--- 1020
QY     366  SerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLys 385
DB     1021  -----GCTTCAAGCCCCACTGTGAATCAACGAGAGCTGTGAACCG 1062
QY     386  ValIleSerLeuThrSerLysThrPylThrValGlyValGlyValValAlaAla 405
      ::::|||||:  ::::|||||:  ::::|||||:
DB     1063  GCTATGAGGCTCTGAAACCA-----ACACGGAGTGTGTTTGTCCCACTG--- 1110
QY     406  ProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAla 425
      ::::|||||:  ::::|||||:  ::::|||||:
DB     1111  -----AAGAAGCTGGCAGACTGCTGCTCAAAATAAGAC 1143
QY     426  GlnPheGlnLysGluValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThr 445
      ::::|||||:  ::::|||||:  ::::|||||:
DB     1144  TCTCAGAAAGATGTGTCTGTGAATGAGGTGTAAGTCTG-----GTATATTGGGCACT 1197
QY     446  GlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGluSerAsnGlnMet 465
      ::::|||||:  ::::|||||:  ::::|||||:
DB     1198  GG-CTGTGAAAAGCAAAAGCAAAAGCCGTGCGATCGAGAGAACTGTAAGATT 1256

RESULT 38
AY103948      1792 bp      mRNA      linear      HTC 25-MAY-2002
LOCUS        Zea mays PC0124731 mRNA sequence.
DEFINITION   Zea mays PC0124731 mRNA sequence.
ACCESSION    AY103948
VERSION      AY103948.1 GI:21207026
KEYWORDS     HTC.
SOURCE       Zea mays.
ORGANISM     Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 1792)
AUTHORS      Hainey C.F., Dolan M., Miao G.H., Vogel J.M., Whitesitt M.S.,
              Arthur L.W., Hanafey M., Morgante M. and Tingey S.V.
              Maize Mapping Project/Dupont Consensus Sequences for Design of
              Overgo Probes
              Unpublished (2002)
              Coe E.C.
              Direct Submission
              Submitted (25-APR-2002) Maize Mapping Project, University of
              Missouri, Columbia, MO 65211, USA
              Location/Qualifiers
              1..1792
              /organism="Zea mays"
              /db_xref="Ensembl:637267"
              /db_xref="taxon:4577"
              /clone="PC0124731"
              /clone_lib="Maize Mapping Project/Dupont Consensus

```



Oy	106	AlaAlaThrThrAlaSerAsnThrAlaIleThrValLeuIleMetGlnThrSerIleGlu	125
Db	191		
Oy	126	AlaSerIysSerMetGluSerThrIleuGlnSerIleGlnSerIleSerAlaIleGlnMet	145
Db	251		
Oy	146	LysCltuVal-GluAlaValValAlaAlaLeuSerGlyLysSer-----SerGlySe	163
Db	305	GTCACtATCCCAAAATCTAAATAATCTCTTGTGCACACAAAACCAGATGCTACAAG	364
Oy	163	rAlaIysIeuGlnThrProGlu-----LeuProIysProGlyValAlaThrProArgSerGl	181
Db	365	AGCCACGGTTACAGACCTCATACTGTTTGTTGGGAGCCAAGATTGACAGATTTCTTCA	424
Oy	181	uValIleGluIleGlyLeuAlaLeuAlaLeuAlaLysAlaIleGlnThrIleuGlnGluAlaThrLy	201
Db	425	G-----CAACACAGCTTAGCACGCGCGAGAGAATTCAAAGTTCAAAGGTGAAGCTGTTTC	478
Oy	201	sSerAlaLeuSerAsn-----TyrAlaSerThrGlnAlaGlnAlaSpGlnTh	217
Db	479	AAACATTTCAGAAAAAACCTCAGACTCCAAACCGTCCAAAGAGAGAGTGAAGAAGAGAGGT	538
Oy	217	rAsnIysIeuGlnGlyLeuGlnAlaIleLysIleLysPlysGluAlaArgGlnGluTrpGcl	237
Db	539	TGATGAGACGGGTGGTGAAGTTAAGACATGAACTGGTCATGTCGCAAGCAACGTAATC	598
Oy	237	nGluMetLysAlaIleGluGlnIlySerLysAspIeuGlnGlyThrMetAspThrValas	257
Db	599	AAGACCAAGGCTGTTCACGCCCTCAAAAACAC-----AGTAATGATATTGTAA	649
Oy	257	nThrValMetIleAlaValSerValAlaIle	267
Db	650	TGCTATTATGGAATTAAACCAATGTAACATATT	680
RESULT 37			
LOCUS	AKO14874		
DEFINITION	AKO14874	1590 bp	mRNA
VERSION	AKO14874.1	GI:12852970	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone.lib.RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Euhaylola; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Mech. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
JOURNAL	Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
MEDLINE	Normalization and subtraction of cap-trapper-selected cDNAs to		
PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes		
REFERENCE	Genome Res. 10 (10), 1617-1630 (2000)		
AUTHORS	3		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,		
JOURNAL	Kono, H., Akiyama, J., Nishik, K., Kitsunai, T., Tashiro, H., Itoh, M.,		
MEDLINE	Sunji, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
PUBMED	Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K.,		
REFERENCE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,		
AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
JOURNAL	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
MEDLINE	11042159		
PUBMED	20499374		
REFERENCE	11042159		

**TITLE**

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913

PUBMED 11076861

**AUTHORS**

Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishik,K., Kiyoasa,H., Kondo,S., Yamanaoka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasutake,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavola,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiya,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staib,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boiffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bolt,C., Fletcher,C., Fujita,M., Gariboldi,M., Gusticich,S., Hill,D., Hofman,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons.P., Machionni,L., Mashima,J., Mazzatelli,V., Mombaerts,P., Nordone,P., Ring,B., Ringuwald,M., Rodriguez,I., Sakamoto.N., Sasaki,H., Sato,K., Schonbach.C., Seya.T., Shibata.Y., Storch.K.F., Suzuki,H., Toyok-Oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilmng.L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsuki.S. and Hayashizaki,Y.

**TITLE**

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE 21085660

PUBMED 11217851

**REFERENCE**

5 (bases 1 to 1590)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Ara,i,A., Atakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bolt,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagati,T., Hara,A., Hayatsu,N., Hill.D., Hiramoto,K., Hiwaka,T., Horii,F., Hume,D., Imotoani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,T., Koijma,Y., Kono,H., Kouda,M., Koya,S., Kutihara,C., Matsuyama,T., Miyasaki,A., Nishi,K., Nomura,K., Numataki.R., Onno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Salto,H., Saito,R., Sakai,C., Sakai,K., Sanjo,H., Sasaki,D., Scribnal,Y., Shibata,K., Shibuta,Y., Shingawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takashihi,F., Tanaka,T., Tejima,Y., Toya,T., Yamanura,T., Yamanaka.I., Yasunishi,A., Yoshida,K., Yoshino,M., Muranatsu.M. and Hayashizaki,Y.

**TITLE**

Direct Submission  
Submitted (10-JULY-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohoma institute; 1-7-22 Suehiro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
[URL:http://genome.gsc.riken-go.jp/](http://genome.gsc.riken-go.jp/), Tel:81-45-503-9222, Fax:81-45-503-9216)

**COMMENT**

Please visit our web site (<http://genome.gsc.riken-go.jp/p/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN. Division of Experimental Animal Research at Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAAGACGAATCCAGACTCTTTTITTTTTTNN 3'], cDNA was prepared by using trehalose thermo activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAAGAAGATTTCAGTAATAAATTAATCAATCCCCCCC 3']. cDNA was cleaved with BamH and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pluscript KS(+) after end excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.  
Location/Qualifiers  
. l\_1590  
"/organism=Mus musculus"  
"/strain=C57BL/6J"  
"/db\_xref=FANTOM\_DB:49215111i16"

**FEATURES**

SOURCE



LOCUS	BG676639		805 bp	mRNA	linear	EST 01-MAY-2001
DEFINITION	602623134F1 NCL_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748097 5'					
ACCESSION	BG676639					
VERSION	BG676639.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.					
TITLE	1 (bases 1 to 805)					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: James Cleaver, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLAM10599 row: c column: 10 High quality sequence stop: 801.					
FEATURES	Location/Qualifiers					
source	1..805 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4748097" /clone_1lb="NCL_CGAP_Skn4" /tissue_type="squamous cell carcinoma" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: This is a NCL_CGAP Library."					
BASE COUNT	182 a 195 c 290 g 138 t					
ORIGIN						
Alignment Scores:						
Pred. No.:	0.0165	Length:	805			
Score:	127.50	Matches:	62			
Percent Similarity:	42.75%	Conservative:	47			
Best Local Similarity:	24.31%	Mismatches:	102			
Query Match:	5.38%	Indels:	44			
DB:	12	Gaps:	8			
US-09-889-314-2 (1-496) x BG676639 (1-805)						
Oy	124 GUGUUAASerLYSerMetGLuSerThrLeuGLuSerLeuGlnSerLeuSerAlaAla	143				
Dd	:::					
Dd	9 GAGGATATAGCCCAAGAAGCAAAGCTGAGGCCGACTCCTGTAC-----	53				
Oy	144 GInMeLySGluValGluAlaValAlaAlaAlaLeuSerGlyLySerSerGlySer	163				
Dd	:::					
Dd	54 CAGACACAAGTAGAAGAGCTGCAGTCACCTCGCGCACAGCATGGGATGTGAAGAAT	113				
Oy	164 AAlaLySLeuGluThrProGluLeuProLySPROGlyValThrProArgSerGluValIle	183				
Dd	:::     :::					
Dd	114 TCMAAGATATGAATAATTCTGAGCTGAACTGTGATCCACAGACTTGATGATCTGAAGCGAC	173				
Oy	184 GIUIeGlyLeuAlaLeuAlaLysAlaAlieGlnThrLeuGlyGluAlaThrLySSerAla	203				
Dd	:::     :::					
Dd	174 ATATGTCAGAACACACATCTCCAATTCGACGACAGCTCATAGATGATCAAGACAG-----	227				
Oy	204 LeuSerAsnTyraLaserThrGlnAlaGlnAlaAspGlnThrAsnLySLeuGlyLeuGlu	223				
Dd	:::					
Dd	228 -----CGTGGCGAGAAATCCCTCAAGAGATGCCAAGAACAAAGCTTAATGACCTG	275				
Oy	224 TysGlnAlaAlaIleLySLeasplySGU-----ArgGlnGluTyrrGlnGlu	238				
	:::     :::					

Db	276	GAGATGCGCCCTGCAGCAGCAGCCACGAAGAAGACTGGGCCGCCCTGTGCGTGAATTACCAGAG	335
Oy	239	MethylsalalagluGlInLysSerLysAspLeuGlUylThrMetAspPThValAsnThr	258
Db	336	CTGATGAACACC---AAGCTGACCCCTGGATCTGGAG-----ATTGGCACCTACAGAAC	386
Oy	259	ValMet-----IleAlaValSerValAlaIleThr	268
Db	387	CTCCCTGGAGGAGAAAGAAAGAGCATGTCGTGAGAAATGTGCCCGAAGCTGATGTCTCT	446
Oy	269	ValliserilValAlaIleAlaIlePheThrcysglAla-----	281
Db	447	GTGACACCAACCCACACCAACCATCATGCTGAGTGAGTGGCACCCAGAGAGTGGCGCGTGGC	506
Oy	282	---GlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValaAlaAlaAlaGlyLy-Al	300
Db	507	TACGCGCTGTGAGTAGACAGCATATGCGTCCGAGAGTGTACTATGCTTGTGAGAGTGGC	566
Oy	300	aAlaGlyAlaAlaAlaAla-----ThrThrValAlaThrGlnIleThr	314
Db	567	GCGCGCGCGCGCGCTGCAGCATAGGCTCCGAGGTGCGACCTATGCTGTGAGAGTGGC	626
Oy	314	rValGlnAlaValAlaGlnAlaValaValLysGlnAlaValIleThrAlaValArgGlnAlaI	334
Db	627	GCGCGCGCGCATGACAGCTACGCGCTCCGGAAGCAGACATGGGGGCTTCACAGAGTGGCTCT	686
Oy	334	eThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAla	348
Db	687	GGAGCGCGCGCGCGCGCAGCTCTGCGCGCGGCGCTGGCGG	729
RESULT 35	B0838510/c	B0838510	644 bp mRNA linear EST 08-Aug-2002
LOCUS	DEFINITION	WHE2911.D01.H01ZS Wheat aluminum-stressed root tip cDNA library	
VERSION	KEYWORDS	Triticum aestivum cDNA clone WHE2911_D01_H01, mRNA sequence.	
SOURCE	ORGANISM	B0838510.1 GI:22142828 Estr. bread wheat. Triticum aestivum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum. 1 (bases 1 to 644) Anderson,O.D., Chao,S., Chin,A., Close,T.J., Gustafson,J.P., Iazo G.R., Rauscher,C.J., Ross,R., Seaton,C.L. and Wilson.C. The structure and function of the expressed portion of the wheat genomes - Aluminum-stressed root tip cDNA library unpublished (2001)	
JOURNAL COMMENT		Contact: Olin Anderson US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel: 5105595773 Fax: 5105595818 Email: oanderson@pw.usda.gov Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: SK primer.	
FEATURES		Location/Qualifiers	
SOURCE		1..644	
		/organism="Triticum aestivum"	
		/cultivar="BH1146"	
		/db_xref="taxon:4565"	
		/clone="WHE2911.D01.H01"	
		/clone_lib="wheat aluminum-stressed root tip cDNA library"	
		/russ_type="Root tip at 1.0 to 1.5 mm stage"	
		/dev_stage="Seedling"	
		/lab_host="E. coli SOLR"	
		/note="Vector: Lambda Uni-ZAP XR, excised phagemid: site_1: EcoRI. site_2: XhoI. Plants were grown under hydroponic conditions, root tips were excised and snap frozen, total RNA was prepared at University of	



```

/dev_stage="Adult"
/lab_host="DH10b"
/note="Vector: pSPoriL; Site_1: SalI; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. CDNA inserts >500 bp
cloned directionally into pSPori 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT      305 a      234 c      135 g      37 t
ORIGIN
Alignment Scores:
Pred. No.:      0.0124      Length:      711
Score:          128.00      Matches:      57
Percent Similarity: 40.34%      Conservative: 37
Best Local Similarity: 24.46%      Mismatches: 107
Query Match:    5.40%      Indels:      32
DB:             13      Gaps:      6

```

```

US-09-889-314-2 (1-496) x BM588321 (1-711)
OY      259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThr 278
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        617 GTTGTGTTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
OY      279 CysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaGly 298
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        557 GTTGTGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
OY      299 GlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 318
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        497 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
OY      319 ValGlnAlaValAlaGlyGlnAlaAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIle 338
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        443 GCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 384
OY      339 LysAlaAlaValIleSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAla 358
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        383 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333
OY      359 LysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyIleThrIleMetIle 376
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        332 GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 273
OY      377 AlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValIleThr 396
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        272 ATTGTTGATTTGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 228
OY      397 ValGlyValAlaGlyValAlaValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGln 416
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        227 TTTCCTGGACAGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 198
OY      417 LeuSerGlnMetGlnAlaGlnAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAla 436
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        197 -----GATGCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 156
OY      437 AlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaAlaSerLysIleAlaSer 456
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        155 TCTGTCGATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 108
OY      457 LysGlnThrGlyLysLeuSerGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIle 476
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        107 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 48
OY      477 LeuLysAlaThrAlaAlaAlaIleSerGlyAlaAlaIleAlaGly 489
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        47 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9
DB      RESULT 32
      BE585195/c      BE585195      773 bp      mRNA      linear      EST 16-AUG-2000
      LOCUS      9-9H-ZO PsajaeZO Phytophthora sojae cDNA, mRNA sequence.

```

```

ACCESSION      BE585195
VERSION        BE585195.1
KEYWORDS       GI:9836144
SOURCE         EST.
ORGANISM       Phytophthora sojae.
                Phytophthora sojae
                Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
                Phytophthora.
REFERENCE      1 (bases 1 to 773)
AUTHORS       Qutob,D., Hraber,P.T., Sobral,B.W.S. and Gijzen,M.
TITLE         Comparative analysis of expressed sequences in Phytophthora sojae
JOURNAL       Plant Physiol. 123 (1), 243-254 (2000)
MEDLINE       20267956
COMMENT       Contact: Gijzen M
                Agriculture and Agri-Food Canada
                1391 Sandford Street, London, Ontario, Canada N5V 4T3
                Tel: 519 457 1470
                Fax: 519 457 3997
                Email: gijzenm@agr.ca.
FEATURES       Location/Qualifiers
                1..773
                /organism="Phytophthora sojae"
                /strain="race 2, strain P6497"
                /db_xref="taxon:6753"
                /clone_lib="PsojaeZO"
                /dev_stage="zoospores"
                /lab_host="E. coli strain XL0LR"
                /note="Vector: pRK-CMV, Site_1: EcoRI; Site_2: XhoI; This
                cDNA library was constructed from polyA+ enriched mRNA
                from zoospores grown in liquid medium. Zoospores were
                released into water and collected by centrifugation at
                2,000g; zoospore-bearing sporangia were induced to
                develop on 5 to 7 d old mycelium colonies grown on V8
                agar by repeated flooding with water. Complementary DNA
                was synthesized from mRNA using an XhoI-poly(dT)
                linker-primer. EcoRI adapters were ligated to the
                blunt-ended cDNA fragments and the products were digested
                with XhoI for directional cloning into lambda ZAP Express
                vector. This lambda library was amplified once using E.
                coli host strain XL1 Blue MRF+. Inserts were then
                subcloned by mass excision using ExSist helper plasmid
                for conversion into phagemid vector pRK-CMV in E. coli
                host strain XL0LR. Sequenced using 73 primer: 5' ATT AAC
                CCT CAC TAA AGG GA 3'."

```

```

BASE COUNT      203 a      266 c      201 g      101 t      2 others
ORIGIN
Alignment Scores:
Pred. No.:      0.0139      Length:      773
Score:          128.00      Matches:      72
Percent Similarity: 31.30%      Conservative: 36
Best Local Similarity: 20.87%      Mismatches: 115
Query Match:    5.40%      Indels:      122
DB:             10      Gaps:      9
US-09-889-314-2 (1-496) x BE585195 (1-773)
OY      69 LysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaGlyLys 88
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        675 AGACATCCGTCGACTGACGACTAGTGCGGACGACGACCTCCGTTGAAGTCCCAATC 616
OY      89 GluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThr 108
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        615 GTTGTACCGGCGGCTGCTGTCGCGGAGCGGACGTCGACGTCGTCGTCGTCGTCGTCGTCGTC 556
OY      109 ThrAlaSerAsnThr---AlaThrLysIleAlaMetGlnThrSerIleGluAlaSer 127
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        555 GGTGCGTCTGTCACGACGCCGCGGCTTGTAGCTGACGACATCCGTCGACGACTTCA 496
OY      128 LysSerMetGluSerThrLeuGlnSerLeuGlnSerAlaAlaGlnMetLysGlu 147
        |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
        495 GATGTGCGTGGCGA-----AATTCCGTCGAGATTGCAATTCGTTGCTACTGCTGCTTCT 442
OY      148 ValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlu 167

```

Oy 100 ThrcgylalserglYAlAAlAAlAThThrThla---SeranThrThrlThrylIAAlA 118  
           |||         |||||         |||         |||         |||  
 Db 248 ---GGGACCCCTGGGAAAGGGGCTACCACCCCTGGGAAGCAGGGCGTGTGCC 304  
 Oy MetcInThrSerIleGlUglUAlAserLySerMetGluSerThrLeuGln 138  
           |||||         :::         :::         |||||         |||  
 Db 305 TCCCAAGCACCAAGCAGGAACCCAGAGAGACTCATAGACACAGACAGAGAGAG---TCA 361  
 Oy SerLeuSerAlaAlaGlnMetLySGLUValJslUAlAValValAlAAlAAlaLeuSerGly 158  
           |||         :::         :::         :::|||||         ||| |||||  
 Db 362 TCCTGCACGTGAGGAGAACCCAGCTCCCAAAGGCCCTTCAGCGCAAGCCCTCAGA 421  
 Oy LysserSer-----GlySerAlaLySLeuGluThrProGluLeuProLyProGlyVal 176  
           |||||         |||||         |||         |||         |||         |||  
 Db 422 AAAACCTCTCAGAGTCGAGCTGCTCCAGCCCCCTGCCAAGAGACTCCCCAGCAAAAGACT 481  
 Oy ThrProCysSerGluValIlleGluIleGlyLeuAlaLeuAlaAlaAlaIleGlnThrLeu 196  
           |||         :::         |||         |||||         |||         |||  
 Db 482 GCCCA---GCCGCCCTGGGAAGACAGAGGCGCTGCAGTTGCCAAGGCCCAGCGGGAG 538  
 Oy GlyAluAlaThrLySerSerAlaLeuSerAnTyrrAlaSerThrGln-----Ala 212  
           |||         :::         |||         :::         :::         |||  
 Db 539 CGGAGGAGAGACTCCAGACAGCAGCGAGATCGACAGTAGAGAGAGAGCGCTGCT 598  
 Oy GlnAlaAspGlnThrAsnLySLeuLy----- 221  
           |||||         :::         |||  
 Db 599 CAGGGCAAGCCTTCAGGAGAAAGCCCCCAGCTCAGAGCCCGCTCGGCCCTGCCAAGAG 658  
 Oy -----LeuGluLySGlnAlaIleLySIIeAspLys 231  
           |||         |||         |||||         :::         :::  
 Db 659 TCCCACAGGAAGAGGCGCTGCCCCAGCACCTCCTCAGAAAACAGGCGCTCAGCGGCCAG 718  
 Oy GlnArg-----GluGluTrpGlnGluMetLySAlaAlaGluGlnLySerLysasp 248  
           |||         ::: |||         ::: |||         ::: |||         |||  
 Db 719 TCCAGGTGGGGAACAGAGAGAGACTCACAAGACGACGAGGAGTACAGACAGTAC 778  
 Oy LeuGluGlyThr 252  
           |||||  
 Db 779 AGAAGAGCCT 790

**RESULT 30**  
**FR0006944/C**  
**LOCUS** FR0006944 619 bp DNA linear GSS 02-MAR-1997  
**DEFINITION** F.rubripes GSS sequence, clone 133B16ac7, genomic survey sequence.  
**ACCESSION** Z90754  
**VERSION** Z90754.1 GI:1867968  
**KEYWORDS** GSS; genome survey sequence.  
**SOURCE** Takifugu rubripes.  
**ORGANISM** Takifugu rubripes.

**REFERENCE**  
**AUTHORS** Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umranta,Y., Williams,G. and Brenner,S.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (18-FEB-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridgeshire, CB10 1SB. Email: bionel@phnmp.mrc.ac.uk  
**COMMENT** Vector: pBluescript II KS  
           V-type: phagemid  
           PRIMER: KS  
           DESCR: One pass dye-terminator sequencing of cosmid cloned genomic sequence.

Alignment Scores:	
Pred. No.:	0.0102
Score:	128.00
Percent Similarity:	49.22%
Best Local Similarity:	30.47%
Query Match:	5.40%
DB:	17
Length:	619
Matches:	39
Conservative:	24
Mismatches:	61
Indels:	4
Gaps:	3

[illegible]

US-09-889-314-2 (1-496) x BG696282 (1-751)	Alignment Scores:	Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
119 MetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeuGln	0.0119	128.50	751	58	49	109	38	6
19 GTCAGGCCCGCATGCAGGATATATGCCCAAGAGCAGCAAAAGCTGAGCCGAGCTCTGTAC	42.13%	22.83%	5.42%	12				
139 SerLeuSerAlaAlaGlnMetLysGluValGlnAlaValValAlaAlaLeuSerGly								
79 -----CAAGCAGCAAGTATGAGAGCTGCAGATCATCTGCGCAGACATGGG								
159 LysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrPro								
124 GATACGTGAGCAAAATTCAGAAATTCAGATTCGAGTGAATGTCGATCCAGACAGACTT								
179 ArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlu								
184 AGATGTGAATTCGACAAATGTCAGAACACAGATCTCCAACTTGACAGTCATCAGTATGAT								
199 AlaThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsn								
244 GCAGAGCAG-----CGTGGCAGAAATGCCCTCAAGGATGCCAAGATCCAGAAC								
219 LysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlu								
286 AGCTGATATGACCTGGAGATGGCCCTGACGACGAGCCAAAGAACACTGGCCGCTGCTG								
234 GlnGluThrGlnGlnMetLysAlaAlaGlu-----								
346 CGTGACTACACAGGACCTGATGACACACCAAGCTGGCCCTGATCTGAGATTCGCACCTAC								
244 -----GlnLysSerLysAspLeuGlnLysThrMetAspThrValAsnThrValMet								
406 AGGACCCCTCTGGAGGAGGAAGAATATCAGATGTCGTGAGACATATGCCCCGAAACGTACT								
261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly								
466 GGTGCTGTGAGCACAAGCACACCAACATTCAGTGAGAGTGGCAGCCAGAGGAGGGGGC								
281 AlaGlyLeu-----AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGly								
526 GGTGGCTACGGCTCTGAGAGTGAAGCAGCTATGCTCCGAGAGGTGACTATGTTCTGGA								
299 Gly-AlaAlaGlyAlaAlaAlaAlaAla-----ThrThrValAlaThrGln								
586 GGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG								
312 IleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgI								
646 GGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG								
332 AlaIleThrAlaAlaIleLysAlaAlaValLysSerGly								
706 GGTCTGTGAGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG								

RESULT	29
LOCUS	BMS56446
DEFINITION	BMS56446 1000 bp mRNA linear EST 20-FEB-2007
ACCESSION	AGNCOCURT_6540507 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737735
VERSION	BMS56446
KEYWORDS	BMS56446.1 GI:18797674
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 1000)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rs9abs@emall.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LRAM12748 row: f column: 08 High quality sequence stop: 614. Location/Qualifiers 1..1000 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5737735" /clone_lib="NIH MGC 88" /tissue_type="duodenal adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: small intestine; Vector: pCMV-SPORT6; site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
BASE COUNT	252 a 284 C 355 g 107 t 2 others
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0179 Length: 1000
Score:	128.50 Matches: 62
Percent Similarity:	39.44% Conservative: 50
Best Local Similarity:	21.83% Mismatches: 115
Query Match:	5.42% Indels: 57
DB:	13 Gaps: 11
US-09-889-314-2 (1-496) x BMS56446 (1-1000)	
OY	5 SerIleSerSerSerSerglyProaspangInLysAsnIlMeKerSerglnvalLeuthr 24
Dd	2 TCAGTTCTACTAAGAAGATCTCCA----- 25
OY	25 SerThrProGlnGlyAlaProGlnGlnAspLysLeuSergLysngInLthrLysGlnIle 44
Dd	26 ---GCAAGAAAGCGGGCCCCAGCCCTGGGAAGGTc---GGGATGTGACACCCTCAGATC 79
OY	45 Gln-----GlnThrArgGlnGlyLysAsnThrGlutMetGluSerAspAla 59
Dd	80 AAAGAGGGGGCCCTGCCGCCAGCAAGAGGGCCCAAGAACAGACAAGAGGAGTCAGAGAGT 139
OY	60 ThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysTrpGlnThrAla 79
Dd	140 ACTGAGAGAGGATCTGAAGTAGTGAGGAGGAGGCCCTGCACAGGACAGCAACCGAGTAAG 199
OY	80 ProGlnGlnGlyAlaAlaAlaGlyLysGluSerSergLysSerGlnLysAlaGlyAlaAsp 99
Dd	200 GCCTCTGAATAAAATTCTCCAGGTCAGACGCTGCGCTCACACCCCTGCCAAG----- 247



Oy	187	LenaLaLeuaLaLysAlaIleGlnThrLeuGlyGluAlaThr-----	200
Db	859	TTTGACCGTAGAGGAAGAAACCACGACGATAGAACAACATCCGTGGCCCTTAGAG	918
Oy	201	LysSerAlaLeuSerAsnTyrAlaSerThrClnAlaIn-----AlaAspGln	216
Db	919	AAGACTTCCTTGAGAACCAAAAGCCCTACCGGGAAGATCATGATTGCATGAC	978
Oy	217	ThraSnLysLeuGlyLeuGluLysGlnAlaIleLysIle-----Asp	230
Db	979	CTCAAT-----ATGAAAAGAGGTGATGCTGTTGGTTCTGTAAACGCCGCCAG	1029
Oy	231	LysGluArgGlu-----GluTyrGlnGluMetLysAla	241
Db	1030	AAAAAAGAAATCAACACCCACCAAGAGTGTTGGAGCACACACTACCATTAAGCA	1089
Oy	242	AlaGlnGlnLysSerLysAspLeuGluGlyThrMetLsp-----	254
Db	1090	ATTTTCCCACGCCCAACTTCACTGTGGCGACACACACMACCTTGTGCATGACGTCCA	1149
Oy	255	-----ThValAsnThrValMetIleAlaLeuSerValAlaIleThrValIle	270
Db	1150	GCAACTACCCCTCAGCATGACGCTCTCTCCCTTGACACATGCTGCTGATCAGCAACT	1209
Oy	271	SerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLau-----	286
Db	1210	TCAgTTACAGGACATTCACACACACCTCCACACACACAGCAACCGTGATTGCCACACGG	1269
Oy	287	---AlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaIla	305
Db	1270	CCTCACCTCCCTCCACACATCCAGCCCTCCCTCTGTGAGTCCCTCCCTTTCCTCAGCC	1329
Oy	306	AlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln-----	320
Db	1330	TCCACCTCCGAGCAATCCAGCTGCATGTGACACACACACACACACACTCCACTCCT	1389
Oy	321	-----AlaValLysGlnAlaValIleThrAla-----ValArgGlnAla	333
Db	1390	TTCGTCCTCCCTCTGTGGAGCAACGAGGTGATGATGACATCAGTTTGGAAACAGCA	1449
Oy	334	IleThrAlaAlaIleLysAlaIleValLysSerGlyIleLysAlaPheIleLysThrIeu	353
Db	1450	GCACTGCTGCTCCCTTCAAGAGAGCTGCACAGTTGCCAAGAAAGCCAGTCTCTGCACAG	1509
Oy	354	VallYsaIalIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThr	373
Db	1510	GCACTGCTCAGAGACTAAACCACCAAGCTGATGCAACCTCACAGTTTGGGCTGGAAGCT	1569
Oy	374	GlmMetIleAlaLysAsnPheProLysLeuSerLysValIleSer	388
Db	1570	GCCTTACTCAGTCTGATCCAGGAGGACCCCTGAGCGTGTCTCAGC	1614
RESULT 27			
LOCUS	FR0008252	443 bp	DNA linear GSS 02-MAR-1997
DEFINITION	F. rubripes GSS sequence, clone 190022DB9, genomic survey sequence.		
ACCESSION	Z92062.1	GI:1869276	
VERSION	232062.1	GI:1869276	
KEYWORDS	GSS; genome survey sequence.		
SOURCE	Takifugu rubripes.		
ORGANISM	Takifugu rubripes.		
REFERENCE			
AUTHORS	Eligar,G., Clark,M., Smith,S., Week,S., Warner,S., Umrania,Y., Williams,G. and Brenner,S.		
JOURNAL	Direct Submission Submitted (18-FEB-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridgeshire, CB10 1SB. Email: biohelp@hmp.mrc.ac.uk		
COMMENT	Vector: pBluescript II KS V_type: phagemid		

PRIMER:	KS
DESCR:	One pass dye-terminator sequencing of cosmid cloned genomic sequence.
FEATURES	Location/Qualifiers
source	1..443 /organism="Takifugu rubripes" /db_xref="taxon:31033" /clone="190022B9" /clone_1fb="cosmid 190022"
BASE COUNT	38 a 111 c 144 g 143 t 7 others
ORIGIN	
Alignment Scores:	
Pred. No.:	0.00564 Length: 443
Score:	128.50 Matches: 34
Percent Similarity:	50.57% Conservative: 10
Best Local Similarity:	39.08% Mismatches: 42
Query Match:	5.42% Indels: 1
DB:	17 Gaps: 1
US-09-889-314-2 (1-496) x FR0008252 (1-443)	
Oy	262 AAlaAAlserValaAlaAlIleThrValIIleSerIlleValaAlaAlaAlIlePheThrCysGlyAla 281
Db	123 GCTGTAGCTGGTACGTAGCTGCCTGTACTCT--GCTGCTGCTAGACCTGCTGCCTGTA 179
Oy	282 GlyLeuAlaGlyLeuAlaAlaGlyAlaAlaAlaValaGlyAlaAlaAlaAlaGlyGlyAlaAla 301
Db	180 GCTGTAGCTGCCTGTAGCTGCCTGCTGTACTGCTGCTGTAGCTGCTGTAGCTGCTGTAGCTGCCT 239
Oy	302 GlyAlaAlaAlaAlaAlaAlaThrThyValaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 321
Db	240 GCCTGCTGCTGTAGCTGCCTGCTGTACTGCTGTAGCTGCTGTAGCTGCTGTAGCTGCCTGCT 299
Oy	322 ValIysGlnAlaValaAlaIleThrValaAlaAlaArgGlnAlaAlaIleThrAlaAlaAlaIleLysAlaAla 341
Db	300 GTAAGCTGCTGCTGTATCACTGCTGCTGCTGTGTTCGCGCTGCTGCTGCTGCTGCTGCT 359
Oy	342 ValIysSerGlyIleLysAla 348
Db	360 GTAGCTGCTGCTGTACTGCT 380
RESULT 28	
Bg696282	751 bp mRNA linear EST 07-MAY-2001
LOCUS	BG696282
DEFINITION	602659426F1 NCI_CGAP_Skn3 Homo sapiens CDNA clone IMAGE:4802622 5' ,
ACCESSION	BG696282
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 751)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: James Cleaver, M.D. CDNA Library Preparation: Life Technologies, Inc. cdna library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LRAM10697 row: C column: 07 High quality sequence stop: 751. location/Qualifiers 1..751 /organism="Homo sapiens"



```

Db 137 CAGCCACCTGCCAGTGAAGAGAGACAGACAGTCGAGCAG-----GCTGATGATCT 190
Qy GluSerSerGluSerGluValAlaGlyAlaAspThr-----GlyVal 102
Db 191 ATCCAAACAGCAGCATGATGACCGGCTCTGCTCCCTCCCTTACGATAGTCTGTA 250
Qy 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 251 GCTGCCGCCGATCCACTCTCCAGACACACTGGGGGCCACTGCTGCCGACAGACA 310
Qy 123 IleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSer 142
Db 311 GTTAACTGTAACACAGCTCCACACCACTCCCGGACCGCTGCTGCCAGACTTCTCT 370
Qy 143 AlaGluMetLysGluValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 162
Db 371 -----ATGACCGAACCACTCTCTCAAGTTGGTCCA-----CAGGGA 406
Qy 163 SerAlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluVal 182
Db 407 TCT-----GAGCGCTTTCTTGTCCAGAGCAGTCAGATGCTCCGTTTCCACAGAT 457
Qy 183 IleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSer 202
Db 458 ATT-----CTTCAACCCCGGAGAGACATGACAGACAGACAGACAGACTTTCAGCAG 508
Qy 203 AlaLeuSerAsn-----TyrAlaSerThrGlnAla 212
Db 509 CCGATGATAGTCGAGTCCAGACACACATGATGATGATGATGATGATGATGATGATGATGAT 568
Qy 213 GlnAlaAspGlnThrAsnLysLeuGluLysGlnAlaIleLysIleAspLysGlu 232
Db 569 GAGTCACTCAGAGATGCTTCAGACAGCTGCTGTAACAACATGACAGTA-----GACAGC 619
Qy 233 ArgGluGluTyrGlnGluMetLysAlaIleGluGlnLysSerLysAspLeuGluLysThr 252
Db 620 TTATCTCTTCTCAAGACTTATGCGCAAGAGCAG-----ACT 658
Qy 253 MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIle 272
Db 659 GCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 706
Qy 273 ValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleAlaValAlaVal 292
Db 706 ----- 706
Qy 293 GlyAlaAlaAlaAlaGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 312
Db 707 -----GCCGACGACAGACAGCAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 757
Qy 313 IleThrValGlnAlaValAlaValAlaValLysGlnAlaValIleThrAla 329
Db 758 TTGGCTATCCCACTGCTTACTACAGCAGCTCAGCAAGCAGTATGCGCCGCA 808
RESULT 26
LOCUS BC007388 1745 bp mRNA Linear HTC 12-JUL-2001
DEFINITION Homo sapiens, clone IMAGE:3830953, mRNA.
ACCESSION BC007388
VERSION BC007388.1 GI:14712402
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1745)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

```

```

COMMENT
Contact: MGC help desk
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC/DC/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadensystemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahney, Stephanie Ford, Julia
Greene, Mark Ketterman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IFAL Plate: 24 Row: d Column: 4
This clone has the following problem: frame shifted.

FEATURES
Source
Location/Qualifiers
1..1745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3830953"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 500 a 534 c 370 g 341 t
ORIGIN

Alignment Scores:
Pred. No.: 0.035 Length: 1745
Score: 129.00 Matches: 93
Percent Similarity: 37.47% Conservative: 70
Best Local Similarity: 21.38% Mismatches: 190
Query Match: 5.45% Indels: 82
DB: 11 Gaps: 18

US-09-889-314-2 (1-496) x BC007388 (1-1745)
Qy 22 ValLeuThrSerThrProGlnGlyValProGlnIleAspLysLeuSerGly---AsnGlu 40
Db 352 ATCATCTCAGACAGCCCGCAGGCG-----CAGCAGGGCTCTCTCCAGCGCAAAATCTT 405
Qy 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
Db 406 CTAACGCAACTCACTCACTCAACCAAGCCAACTCTTACAGTCGACGCAAGCATCAC 465
Qy 61 IleAlaGly---AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAla 79
Db 466 CTCACCTCCACGACCAACCCCAACAGCAGCAATATGACGACCAATTCAGACACTT 525
Qy 80 ProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAsp 99
Db 526 CCACAGAGCAGTCAACACCAAGCGAATGTACTCCAGCTTGAGAGCGCCAGTGCAC 585
Qy 586 CTTGAGAGCTTGAGCAGCTTCCCAAGACCTTCAACCAAGACGAATCAAACTGATTTC 645
Db 120 ---GlnThrSerIle-----GluGluAlaSerLysSer 129
Qy 646 ACTCAGGGTATGTTGGCTCGCTATGGGGAATATATGAAATGATTCAGCCCAACT 705
Qy 130 MetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAlaGluMetLysGluValGlu 149
Db 706 ACCATCTCTCGATTGAAGCCTTG---AACCTCAGCTTTAAGACATGTGCAAGTTGAAG 762
Qy 150 AlaValAlaValAlaAlaLeuSerGlyLysSer-----SerGlySerAlaLysLeuGlu 167
Db 763 CCACCTTTAGAGAACTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 822
Qy 168 ThrProGlu---LeuProLysProGlyValThrProArgSerGluValIleGluIleGly 186
Db 823 AGCCCAAGTGCCCTGAATTCACAGGAATTCAG-----GGC 858

```



TITLE Genomic exploration of the hemiascomycetous yeasts: 11.  
 JOURNAL Kluveromyces lactis  
 MEDLINE FEMS Lett. 487 (1), 66-70 (2000)  
 PUBMED 20584721  
 11152886  
 3 (bases 1 to 892)  
 REFERENCE Genoscope.  
 AUTHORS Direct Submission  
 TITLE  
 JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT This GSS is part of a random genomic sequencing program of thirteen  
 yeast species: *Saccharomyces bayanus* var. *uvareum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluveromyces thermotolerans*, *Kluveromyces*  
*lactis* var. *lactis*, *Kluveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
 5 kb were prepared and both extremities were sequenced. See  
 keywords for description of this sequence and for the sequence of  
 the other extremity of this insert.

FEATURES  
 source  
 1. 892  
 /organism="Kluveromyces lactis"  
 /strain="CLIB 210"  
 /variety="lactis"  
 /db\_xref="taxon:28985"  
 /clone="BA0AB037D05"  
 /clone\_1id="BA0AB"  
 /clone\_2id="BA0AB"

BASE COUNT 274 a 275 c 120 g 222 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.00854 Length: 892  
 Score: 131.00 Matches: 59  
 Percent Similarity: 36.52% Conservative: 48  
 Best Local Similarity: 20.14% Mismatches: 160  
 Query Match: 5.53% Indels: 26  
 DB: 17 Gaps: 6

US-09-889-314-2 (1-496) x CNS075P4 (1-892)

Oy 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27  
 |||||  
 |||||  
 Db 45 TCCAGCACTGATTCCTTCAGAACACCTTCAGCACTCTTCGAGAGAACCTTCACCACT 104  
 |||||  
 |||||  
 Oy 28 GlnGlyValProGlnGlnAspLysLeuSerGlnGluThrLysGlnIleGlnGlnThr 47  
 |||||  
 |||||  
 Db 105 GATTCCTTCGAAACACTTCCTCCGAGCAACTTCGAGAACCTTCACCACTGATTCCTCT 164  
 |||||  
 |||||  
 Oy 48 ArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLys 67  
 |||||  
 |||||  
 Db 165 GAACACCTTCAGCACTTCCTTCGAGAACACTTCACCACTGATTCCTTCGAGAACCT 224  
 |||||  
 |||||  
 Oy 68 AspLysThrSerSer-----ThrLysThrSerGlnThrAlaProGlnGlnGlyAla 85  
 |||||  
 |||||  
 Db 225 TCAACCACTTCCTTCGAGAGAACCTTCACCACTGATTCCTTCGAGAACCTTCACCACT 284  
 |||||  
 |||||  
 Oy 86 AlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAla 105  
 |||||  
 |||||  
 Db 285 TCTTCGAGAGAACCTTCGAGAACCTTCGAGAACCTTCGAGAACCTTCGAGAACCT 344  
 |||||  
 |||||  
 Oy 106 AlaAlaThrThrAlaSerAsnThrAlaThrIleAlaMetGlnThrSerIleGlnGlu 125  
 |||||  
 |||||  
 Db 345 GAAACTTCAGCACTTCCTTCGAGAACCTTCGAGAACCTTCGAGAACCTTCGAGAACCT 404  
 |||||  
 |||||  
 Oy 126 AlaSerLysSerMetGlu-----SerThrLeuGluSerLeuSerLeuSerAlaAla 143  
 |||||  
 |||||  
 Db 405 TCAACCACTTCCTTCGAGAGAACCTTCACCACTGATTCCTTCGAGAACCTTCACCACT 464  
 |||||  
 |||||  
 Oy 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163  
 |||||  
 |||||  
 Db 465 TCTTCGAGAGAACCT-----TCAACCACTGATTCCTTCGAGAACCTTCGAGAACCT 500  
 |||||  
 |||||

Oy 164 AlaLysLeuGlnIleThrProGlnLeuProLysProGlnValThrProArgSerGluValIle 183  
 |||||  
 |||||  
 Db 501 CAGCCTTCGATTCAGCACTTCGAGAACCTTCACCACTGATTCCTTCGAGAACCTTCACCACT 557  
 |||||  
 |||||  
 Oy 184 GluIleGlyLeuAlaLeuAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203  
 |||||  
 |||||  
 Db 558 -----TCAACCACTGATTCCTTCGAGAACCTTCGAGAACCTTCGAGAACCT 593  
 |||||  
 |||||  
 Oy 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223  
 |||||  
 |||||  
 Db 594 TCTTCGAGAGAACCTTCACCACTGATTCCTTCGAGAACCTTCGAGAACCTTCGAGAACCT 653  
 |||||  
 |||||  
 Oy 224 LysGlnAlaIleLysIleAspLysGluArgGlnGluThrGlnGlnMetLysAlaAlaGlu 243  
 |||||  
 |||||  
 Db 654 GAAACCACTTCAGCACTTCATTCACCACTGATTCCTTCGAGAACCTTCGAGAACCTTCACCA 710  
 |||||  
 |||||  
 Oy 244 Gln-----LysSerLysAspLeuGlnGluThrMetAspThrValAsnThrVal 259  
 |||||  
 |||||  
 Db 711 CAACCTTCAGCACTGATTCATTCAGAACCACTTCACCACTTCGAGAACCTTCGAGAACCT 770  
 |||||  
 |||||  
 Oy 260 MetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCys 279  
 |||||  
 |||||  
 Db 771 ACTATTGCTGACTACCGCTTACCACTGATTCATTCGAGAACCTTCGAGAACCTTCACCACT 830  
 |||||  
 |||||  
 Oy 280 GlyAlaGlyLeuAlaGlyLeuAlaIleGlnThrValAlaAlaVal 292  
 |||||  
 |||||  
 Db 831 ACTGTCCATTTGATTCACCACTGATTCATTCGAGAACCTTCGAGAACCTTCGAGAACCT 869  
 |||||  
 |||||

RESULT 24  
 CNS05HLZ 1101 bp DNA linear GSS 26-MAY-2000  
 LOCUS CNS05HLZ  
 DEFINITION Tetradodon nigroviridis genome survey sequence T3 end of clone  
 010E07 of library A from Tetradodon nigroviridis, genomic survey  
 sequence.  
 ACCESSION AL337760.1 GI:8231518  
 VERSION AL337760.1  
 KEYWORDS GSS: genome survey sequence.  
 SOURCE Tetradodon nigroviridis.  
 ORGANISM Tetradodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.  
 1 (bases 1 to 1101)  
 2 (bases 1 to 1101)  
 3 (bases 1 to 1101)  
 Rost-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,  
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
 Saurin,W. and Weissenbach,J.  
 Human gene number estimate provided by genome wide analysis using  
 Tetradodon nigroviridis DNA sequence  
 Unpublished

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

COMMENT  
 JOURNAL  
 REFERENCE  
 AUTHORS

FEATURES  
 source  
 1. 1101  
 /organism="Tetradodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="010E07"  
 /clone\_1id="A"  
 /clone\_2id="A"  
 /clone\_3id="A"  
 location/Qualifiers  
 1. 1101

This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetradodon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/Tetradodon.

DEFINITION T7 end of clone BB0A002G01 of library BB0A from strain CBS 4732  
 of *Pichia angusta*, genomic survey sequence.  
 ACCESSION AL431231  
 VERSION AL431231.1 GI:12214643  
 KEYWORDS GSS:  
 SOURCE *Pichia angusta*.  
 ORGANISM *Pichia angusta*  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; *Pichia*.  
 REFERENCE 1 (bases 1 to 888)  
 AUTHORS Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
 Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potter,S.,  
 Saurin,W., Tekaiia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
 Wincker,P. and Weissenbach,J.  
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
 yeast species for molecular evolution studies  
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
 MEDLINE 20584711  
 PUBMED 11152876  
 REFERENCE 2 (bases 1 to 888)  
 AUTHORS Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.  
 and Dujon,B.  
 TITLE Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia*  
*angusta*  
 JOURNAL FEBS Lett. 487 (1), 76-81 (2000)  
 MEDLINE 20584723  
 PUBMED 11152888  
 REFERENCE 3 (bases 1 to 888)  
 AUTHORS Direct Submission  
 TITLE Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,  
 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :  
 seq@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT This GSS is part of a random genomic sequencing program of thirteen  
 yeast species: *Saccharomyces bayanus* var. *uvaurum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
 5 kb were prepared and both extremities were sequenced. See  
 keywords for description of this sequence and for the sequence of  
 the other extremity of this insert.

FEATURES  
 source  
 1..888  
 location/Qualifiers  
 /organism="Pichia angusta"  
 /strain="CBS 4732"  
 /db\_xref="taxon:4905"  
 /clone="BB0A002G01"  
 /clone\_1lb="BB0A"  
 /note="end : 77"

BASE COUNT 225 a 190 c 273 g 194 t 6 others

ORIGIN

Alignment Scores:  
 Pred. No.: 0.00674 Length: 888  
 Score: 132.00 Matches: 58  
 Percent Similarity: 38.82% Conservative: 41  
 Best Local Similarity: 22.75% Mismatches: 122  
 Query Match: 5.57% Indels: 34  
 DB: 17 Gaps: 6

US-09-889-314-2 (1-496) x CNS076CP (1-888)

QY 5 SerIleSerSerSerGergIProAspAsnGlnIlyAsnIleMetSerGlnValIleThr 24  
 DB 813 TCTGCACAAAGCTCATTTTGGCGAAGCCTTTTTCAGTGCCTGGAAACACTTTGCCACAGCC 754

QY 25 SerThrProGlnGlyValProGlnGlnAspIlySerLeuSerGlyAsnGlnIlyThrIlyGlnIle 44  
 DB 753 TCCCGTAGCGGGGAGATATCTCAATTCAGCAGCAGCAGACATGCTCCACTGCTTAGATAT 694

QY 45 -----GlnGlnThrArgGlnGlnIlyAsnThrGluMetGluSerAspAlaThrIleAla 62  
 DB 693 AATGCCAGCTCCTCAGTCGATCTATCTACGATACAGGGGAGCATCCAGTTGGCATCATCTTC 634

QY 63 GlyAlaSerGlyLysAspIlyThrSerThrThrIlyThrIlyThrIlyAlaProGlnGln 82  
 DB 633 GGCTCCAGCAGATGCTCTTCCTTCAGCATATTCAGCAACGACGACCGCTTCACACTAT 574

QY 83 GlyValAlaIlaIlaGlyIlyGluSerSerGlu-----SerGlnIly 95  
 DB 573 GGCTTAGCCTCTTTCAATATCCACATCAGATACATTCAGGGGCTTTACCGCAACT 514

QY 96 AlaGlyIlaIlaIlyThrGlyValSer-----GlyAlaIlaIlaIlyThrIlaSerAsn 112  
 DB 513 GCGGAGCGTCAGATGCAATTTCCAGCCAAACAGACATCCAGCACTTTGTCGAC 454

QY 113 ThrAlaThrIlyIleAlaMetGlnThrSerIleGlnGluIlaIlaSerIlySerMetGluSer 132  
 DB 453 GGTTCAGATCATTTGCTGCTCTAC-----GCTTCTGGCACTGTCCTAGT 406

QY 133 ThrLeuGluSerLeuGlnIlySerLeuSerAlaIlaGln----- 144  
 DB 405 GGAGCTAGTGGAAATTCAGTGTTCCTGCGCAGCTCGAGTATATTTCAGTGTGCC 346

QY 145 -----MetIlyGluValGluAlaValAlaValAlaIlaIlaIleuSerGlyIlySer 160  
 DB 345 CCCAGCTCCATTTTCAACAGCGCTTCACAGCAAGTCCAGCGGTGACTTATCTAGTCT 286

QY 161 SerGlySerAlaIlyIleuGlnIlyThrProGluLeuProIlyProGlyValThrProArgSer 180  
 DB 285 GCCTTAGCGCTTCACAGTCCAGTCCACATTCCTTAAGACTGCTTCGACGCGCTTCTGC 226

QY 181 GluValIleGluIleGlyLeuAlaIleuAlaIleuAlaIleuIlyThr----- 195  
 DB 225 AGCTCCCTCAGCGCAACGTCGCCGCGCAGAGCTTCACAGCAGCAGCAACGGGTC 166

QY 196 ---LeuGlyIlyIlaIlyThrIlySerAlaIleuSerAlaIlyIlaSerThrIlaGlnAla 214  
 DB 165 ATTGTCTCGAGCAACACTGCTTCAGTGTTCATGCTATGTTCCACCGACGCGCAGCGTC 106

QY 215 AspGlnThrAsnIlyLeuGlnIlyGlnIlyGlnAlaIleuIlyIle 229  
 DB 105 ACATCTAGCTCAGGTGCTCTAGCAAAATACAGATTATCTCTG 61

RESULT 23  
 CNS075P4  
 LOCUS 892 bp DNA linear GSS 07-JUL-2001  
 DEFINITION clone BA0AB037D05 of library BA0AB from strain CLIB 210 of  
*Kluyveromyces lactis*, genomic survey sequence.  
 ACCESSION AL430382  
 VERSION AL430382.1 GI:12213576  
 KEYWORDS GSS:  
 SOURCE *Kluyveromyces lactis*.  
 ORGANISM *Kluyveromyces lactis*  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; *Kluyveromyces*.  
 REFERENCE 1 (bases 1 to 892)  
 AUTHORS Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
 Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potter,S.,  
 Saurin,W., Tekaiia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
 Wincker,P. and Weissenbach,J.  
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
 yeast species for molecular evolution studies  
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
 MEDLINE 20584711  
 PUBMED 11152876  
 REFERENCE 2 (bases 1 to 892)  
 AUTHORS Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,  
 Duchateau-Nguyen,G., Lemaire,M., Marmisse,R., Montchocker,R.,  
 Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.

Qy 286 LeuAlaIaIaGlyAlaIaValGly---AlaAlaIaIaIaGlyAlaIaIaGlyAlaIa 304  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 512 GTTCGGCAGAGAGCTCGTTGGGTAGAGAGCTGCAGTAGAGAGCTCGTAGAGAGTGC 453  
Qy 305 AlaAlaIa 321  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 452 GTAGAGCTCGTAGAGAGCTCGTAGAGAGCTCGTAGAGAGCTCGTAGAGAGCTCTTA 393  
Qy 322 ValLysGlnAlaValIleThrAlaValAlaGlnAlaIleThrAlaIleIleLysAlaIa 341  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 392 GTAGAGCTCGTAGAGAGCTCGTAGAGAGCTCGTAGAGAGCTCGTAGAGAGCTCGTAG 333  
Qy 342 ValLysSerGlyIleLysAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 361  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 332 GAAGAAGATGGGATCTCAGAGAGGTTAGAACTGTGGGGCTCAGAGAGGCTCAGAGAGT 273  
Qy 362 SerLysGlyIleSerLysValPheAlaLysGlyIleThrGlnMetIleAlaLysAsnPro 381  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 272 --GAGGTGCTCTCCTCAGTAGAGAGTGGGACACAGAGAGAGAGAGAGAGAGAGAG 216  
Qy 382 LysLeuSerLysValIleSerSerLeuThrSer----- 392  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 215 GAAGCTCTGCTGGAAGTCTCGCTCGGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 156  
Qy 393 -----LysTrp-----ValThrValGlyValGlyValValVal 403  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 155 GACATGCCAGACTGTTAGAGACCGGAGAGACGACGACGACGAGAGAGAGAGAGAGAG 96  
Qy 404 AlaAlaProIaLeuGly 409  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 95 TCAGCAACGAAAGTGGGA 78  
RESULT 21  
AY107616 1027 bp mRNA linear HTC 25-MAY-2002  
LOCUS  
DEFINITION Zea mays PC0064762 mRNA sequence.  
ACCESSION AY107616  
VERSION AY107616.1 GI:21210694  
KEYWORDS HTC.  
SOURCE  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 1027)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1027)  
Coe,E.C.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
Location/Qualifiers  
1. 1027  
/organism="Zea mays"  
/db\_xref="MaizeDB:633478"  
/db\_xref="taxon:4577"  
/clone="PC0064762"  
/clone.lib="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
assigns to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACS in conjunction with the Maize  
Mapping Project"

Alignment Scores:  
Pred. No.: 0.00738 1027  
Score: 132.50 78  
Percent Similarity: 39.80% Matches: 41  
Best Local Similarity: 26.09% Mismatches: 117  
Query Match: 5.59% Indels: 64  
DB: 11 gaps: 12  
US-09-889-314-2 (1-496) x AY107616 (1-1027)  
Qy 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGln 43  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 54 ACAGAAACCCCAAG-----CCCCAAAGGCCCAACTCACACCTCCCAATGGAGACG 107  
Qy 44 IleGlnIleThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 108 GC-GACGAACCTTCGCTCCGCCACGCGCGCGCCGCGAGAGAGAGAGAGAGAGAGAG 166  
Qy 64 AlaSerGlyLysAspLysThrSerSerThrLysThrLysThrLysThrLysThrLys 78  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 167 GTCCGAG 226  
Qy 79 -----AlaProGlnGlnGlyValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 94  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 227 GACGACAAACCCGCTCCTACGCGCGCGCTGTGAAGAGAGAGAG----- 271  
Qy 95 LysAlaGlyAlaAspThrGlyValSerGlyAlaAlaIaIaIaIaIaIaIaIaIaIaIa 114  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 272 -----GGCGGTGCCACACACCGCAACCTCCTACGCA 304  
Qy 115 ThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGlnSerLys 134  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 305 GCT--GTCCGCCGCCCAAGCGAGATCGAGAG----- 334  
Qy 135 GluSerLeuGlnSerLeuSerAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 154  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 335 -----CTCCGCCGCCCAAGCTCGACTCGAGAGAGAGAGAGAGAGAGAGAGAG 379  
Qy 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro 174  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 380 CGCCGTGAGAACAGTCCCTCGCCGAGAG----- 409  
Qy 175 GlyValThrProArgSerGluValIleGluIleGlyLeuAlaIaIaIaIaIaIaIaIa 194  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 410 -----ACGACCGCAGCGAGGAGATC-----TTACGACAGCGCGGAGAGAGCTC 454  
Qy 195 ThrLeuGlyGlu---AlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGln 213  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 455 ACCATCGCGGAGAGTTCGCGCCACCTCTCCGAGAAAGAGAGTGCCTCGCGCGAGAG 514  
Qy 214 AlaAspGlnThrAsnLys---LeuGlyLeuGlnLysGlnAlaIaIaIaIaIaIaIaIaIa 232  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 515 GTCGAACCGCTCGATGCTGTCTCAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 574  
Qy 233 ArgGlnGlu---TyrGlnGlnMetLysAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 250  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 575 AGACACGAGAGAGTCCAG 634  
Qy 251 GlyThrMetAspThrValAsnThrValMetIleAlaIaIaIaIaIaIaIaIaIaIaIa 270  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 635 -----GAGAGATCAGAGAGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685  
Qy 271 SerIleValAlaIa 290  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 686 GAG 742  
Qy 291 AlaValGlyAlaIa 309  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 743 GCGATGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 799  
RESULT 22  
CNS076CF/C CNS076CP 888 bp DNA linear GSS 07-JUL-2001  
LOCUS

Db 358 GSCACACCCMCACAC-----CGGCCAGCCCAACAGCGGCCACAGCCGCCSCC 405

QY 156 uSerGlytysSerSerGlySerAlaIalysLeuGluThrProGluLeuProIlySerProGlyVa 176

Db 406 CCCAGAACACACCGGGGAAACAACAGAG---CCGCCVCCMCAGGCCCGAGCCAGACACA 462

QY 176 lThrProArgSerGluValIleGluIleGlyLeuAlaIalysAlaIleGluThrIle 196

Db 463 CCGMCCAGACACMGAGCCMGAG----- 487

QY 196 uGlyAlaThrIlySerSerAlaLeuSerAsnTyraIalSerThrGluAlaGluAlaIalaspI 216

Db 488 -----ACACAGACACCGMCGACAGACAGACAGACAGAGAMACAGAGACAGCA 537

QY 216 nThrAsnIlyLeuGlyLeuGluIlySerGluAlaIleIlySleAspIlySerGluArgIleGlu 236

Db 538 GGAGCGACGAGACAGACAGACAGAGAGAG---AAACARAGACACAGACAGAGAGAGAGCA 594

QY 236 rGlnIlyMet-----IysAlaIalagIlySerIlySAspIlySerGluIlyThrMetAs 254

Db 595 GCAGGAGCAGACAGARAAAMAGACAGAGAGAGAGCA--GCAGAGCAGCAGAG 653

QY 254 pThrValaIalThrValaIleMetIleAlaValSerValaIalIleThrValaIleSerIleValaI 274

Db 654 AGAACACAGACAGACAGAGAGAGCAGCAGCA-----GC 686

QY 274 aAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIalIalThrThrValaIa 294

Db 687 ACAGAGAGAAACAGACAGACAGAGAGAGCAGAG---GCAGCAGAGACAGACAGAGAAAM 743

QY 294 aAlaIalIalGlyIlyAlaIalGlyAlaIalIalIalIalThrThrValaIa 310

Db 744 AGAGACAGACAGAGAGAGACAGAGACAGAGAGAAACAGAGACAGCA 792

RESULT 20

CNS06RUI/ 1123 bp DNA linear GSS 05-JUL-2001

LOCUS T3 end of clone AM0AA017601 of library AM0A from strain CLIB 89 of

DEFINITION Yarrowia lipolytica, genomic survey sequence.

ACCESSION AL412432

VERSION AL412432.1 GI:12183046

KEYWORDS GSS.

SOURCE Yarrowia lipolytica.

ORGANISM Yarrowia lipolytica.

REFERENCE 1 (bases 1 to 1123)

AUTHORS Soulier,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boloiti-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neveuglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saudin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Winkler,P. and Weissenbach,J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of

JOURNAL Yeast species for molecular evolution studies

MEDLINE FEMS Lett. 487 (1), 3-12 (2000)

PUBMED 20584711

REFERENCE 2 (bases 1 to 1123)

AUTHORS Casaregola,S., Neveuglise,C., Lepingle,A., Bon,E., Feynerol,C., Artiguenave,F., Winkler,P. and Galliardin,C.

TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia

JOURNAL lipolytica

MEDLINE FEMS Lett. 487 (1), 95-100 (2000)

PUBMED 20584727

REFERENCE 3 (bases 1 to 1123)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,

COMMENT 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen

Yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

source

1. 1123

/organism="Yarrowia lipolytica"

/strain="CLIB 89"

/db\_xref="taxon:4952"

/clone="AM0AA017601"

/clone\_1lb="AM0AA"

/note="end : T3"

BASE COUNT 184 a 401 c 240 g 295 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 0.00528 Length: 1123

Score: 134.50 Matches: 82

Percent Similarity: 39.89% Conservative: 64

Best Local Similarity: 22.40% Mismatches: 167

Query Match: 5.68% Indels: 53

DB: 17 Gaps: 13

US-09-889-314-2 (1-496) x CNS06RUI (1-1123)

QY 74 ThrIlyThrGluThrAlaProGlnGlyValAlaIalGlytysGluSerGluSer 93

Db 1106 AGTGAAGACACACAGCGCTCTCTCAGACTGTGTG-----GGACCACCTGGAACACACTCC 1053

QY 94 GlnIlyAlaGlyAlaIalAspThrGlyValSerGlyAlaIalAlaIalThrThrAlaSerAsnThr 113

Db 1052 AGAGARACTCTCTGGAGACACTGCARAGACTCGGTAGACCTCGGAATGTAGGGCTTCACG 993

QY 114 AlaThrIlyIleAlaMetGlnThrSerIleGluGluAlaSerIlySerMetGluSerThr 133

Db 992 GACGAGAG-----GGCTCTCCAGAAACACCGGGA-----GATGACACA 954

QY 134 LeuGluSerLeuGlnSerLeuSerAlaIalGluMetIlySerGluValGluAlaValaIalVal 153

Db 953 GGAGAAACACTCTCAGACAGAGACCTCGCGCTGTGTAACACAGACAGATGACAGCTGTGTG 894

QY 154 AlaIalLeuSerGlytysSerSerGlySerAlaIlySerGluIlyThrProGluLeuProIlys 173

Db 893 GCTGCAGAGTTGACACACTCAGAGGTAGAGTAAAGTCTCAGTGGCGCTGCTGTGCTGA 834

QY 174 ProGlyVal-----ThrProArgSerIlyValIleGluIleGlyLeu 187

Db 833 GGAGGAATGTAGGGCTTCACAGATGAATAATCCATGTGCGGAACACACGTGAAGAAGATATA 774

QY 188 AlaLeuAlaIlyAlaIleGlnThrLeuGlyAlaIalThrIlySerAlaLeuSerAsnTyraI 207

Db 773 GCAGAGACCCAGAGCTGCTCTCTCAGTGGG-----CTGGCAGACAGCTG 732

QY 208 AlaSerThrGlnAlaGlnAlaIalAspGlnThrAsnIlySerGlyLeuGluIlySerGlnAlaIle 227

Db 731 GTCCTCCACATG-----GAGGTAGGAGATGACAGGCACCGCAGCTG 693

QY 228 ---LysIleAspIlySerGluArgIlyGluIlyThrGlnIlyMetIlySerAlaIalGluIlySer 246

Db 692 ACTTCACCTGGAAGAGGACTCGAGAGAGTCTCGGTCTCGAGAGGAGCCAGGAGCCACA 633

QY 247 LysAspLeu---GluGlyThrMetAspThrValaIalThrValaIleAlaIalSerValaI 265

Db 632 GAAGACATACCAGAGAAGAGAAGAAATGTAGTGGTGGAGGCTCCAGAGAGGTG 573

QY 266 AlaIleThrValaIleSerIleValaIalIlePheThrCysGlyAlaGlyLeuAlaGly 285

Db 572 GAAGGCACAGCAGCGAGGTGAGACCTCGTGAAGAGATCAAGGAGCAGCTCCGAA 513





through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 53 Row: 9 Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis

FEATURES  
 This clone has the following problem: incomplete processing.  
 Location/Qualifiers

source

1.2589  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5364153"  
 /issue\_type="Eye, retina, mouse strain C57Bl/6"  
 /clone\_lib="NID\_MGC\_94"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 BASE COUNT 645 a 703 c 615 g 626 t  
 ORIGIN

#### Alignment Scores:

Pred. No.: 0.0154 Length: 2589  
 Score: 135.00 Matches: 93  
 Percent Similarity: 32.368 Conservative: 40  
 Best Local Similarity: 22.638 Mismatches: 155  
 Query Match: 5.70% Indels: 123  
 DB: 11 Gaps: 15

US-09-889-314-2 (1-496) x BC028951 (1-2589)

QY 76 ThGluThrAlaProGlnGlnGly--ValAlaAlaGlySerGlnSerGln 94  
 Db 41 AGTTCACAGACTCCAGAGCCCAATCTCTGCACAGATCAAAAGCTGTGAGCAAGCT 100  
 QY 95 LysAlaGlyAla-----AspThrGlyValSerGlyAlaAlaAlaThrTrpAlaSer 111  
 Db 101 AGGACCGGAACATCATCAATGAGCAAGAAACAGAGATGCTGCACACAGCCGCC 160  
 QY 112 AsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGluAlaSerLysSerMetGlu 131  
 Db 161 ACCACCGCAGGCCA---GCCAGTCTCCCAACAGCAACAGCCGAGTGAAGTACGAC 217  
 QY 132 SerThrLeuGlnSerLeu-GlnSerLeuSerAlaAlaGlnMetLysGluValGluAla 151  
 Db 218 TCCACAGCCCCACAGCAACAGCAG----- 242  
 QY 151 ValValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlu 171  
 Db 243 -----AGTCTCAGCTTACACAGCTCAGCAGC 268  
 QY 171 uProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAla 191  
 Db 269 CCCACGGCTGGGGCTTGTGTCACAGGCACCAAGT6--ACCAACCTTCAGGTGGCCG 325  
 QY 191 salalileGlnThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyrAlaSerThrG1 211  
 Db 326 G-----CTTACCGGAGTCCCTACTCTCTCAAGCTCAGCAGCAAGGAGAGATCA 373  
 QY 211 nalaglnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLys 231  
 Db 374 AACCTCAGAGCCCGCCAGCTGCTCAGGTGGCTGGCAAAACCTCCCGGTG----- 422  
 QY 231 sgluArgGluGlnTyrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGln 251  
 Db 422 ----- 422  
 QY 251 yThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSe 271  
 Db 423 -----GATATGTCGCCAGCAGCTGTGGTCTC 448  
 QY 271 rIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAl 291  
 Db 449 C-----TCGCCAGGGGTCCACAGCCTGCTGCTGAACGTTGC 484  
 QY 291 aValGlyAlaAlaAlaGlyAlaAlaGlyAlaAlaGlyAlaAlaAlaThrThrValAlaThr 311

Db 485 AGGATCAGGGGTGGCAATGTGGACAGCCCAAAAACAGCAGGGCAGCGGTGGGCCA 544  
 QY 311 rGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAr 331  
 Db 545 ACCGTCATGATCCACAGCTGCTGAAGTAAAGACAGCAG-----ACAGCGCTCA 595  
 QY 331 gGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIle 351  
 Db 596 ACACAGCAG----- 602  
 QY 351 sThrLeuValLysAlaIleAlaValAlaLysSerLysGlyIleSerLysValPheAla 371  
 Db 603 -----AAGGCCATTCAGCCACCAAGTGGCCAGGCGCAGGCTGTTCAGCAGAA 652  
 QY 371 s---GlyThrGlnMetIleAlaLysAsnProLysLeuSerLysValIleSerLys 390  
 Db 653 GCTGACCAACCCACAGATCACTACTCAGGGGCCAGC---CAGAAAGCTC----- 698  
 QY 390 uThrSerLysTrpValThrValGlyValAlaValAlaAlaAlaProAlaLeuGly 410  
 Db 699 -----GCTTATGCTGCCACAGCCCTTAAAC 727  
 QY 410 sGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln-----AsnValAlaG1 426  
 Db 728 CCACTTCTCAGCCACACCCATCTCCAGGCCAGCAAACTGCCGCGACTCAGCAGTTCA 787  
 QY 426 nPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrG1 446  
 Db 788 GACTCAGATCAGGATGTCACAAATCTCTCAAGTTGTCACACCAACAAACCTGTGGCCAG 847  
 QY 446 nPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerGlnMet 466  
 Db 848 TATACAAACAAGTGCATCA-----GCTTCTCAGCAGGCTTCTCCACAGAGGTGACGT 901  
 QY 466 rGlnLysAlaThrLysLeuGlyAlaGlnIle 476  
 Db 902 CACACAGCAACGGCAGCAGGCAACAGGTA 932

RESULT 18  
 AA930444 485 bp mRNA linear EST 23-Apr-1998  
 LOCUS vs59603.r1 Strizagene mouse skin (#937313) Mus musculus cDNA clone  
 DEFINITION IMAGE:1150612.5' similar to TR:008884 008884 HIGH-GLYCINE TYROSINE  
 KERATIN TYPE II.4.; mRNA sequence.

ACCESSION  
 AA930444  
 VERSION  
 AA930444.1 GI:3079225  
 KEYWORDS  
 EST.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus

REFERENCE  
 AUTHORS  
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.

TITLE  
 JOURNAL  
 COMMENT  
 The WashU-HMI Mouse EST Project  
 Unpublished (1996)  
 Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 MGI:623820  
 Possible reversed clone: similarity on wrong strand  
 Seg primer: -28m13 rev1 ET from Amersham  
 High quality sequence stop: 216.



```

Db 1502 GACCTAAGCTTT-----GACCTACATCCATCCACCAACCTGAC 1540
Oy 78 ThrAlaProGlnGlnGlyValAlaAlaGlyLysGlySerSerGlySerGlnLysAlaGly 97
Db 1541 TTGGGCCATCCCTGGGA---AAACCTGGTGAACCTCTCCCAAGACAGCTCCAAATCATCT 1597
Oy 98 AlaAspThrGlyValSerGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 117
Db 1598 CCATCTGCTTAAATACATCTTTAGTAAATCCACTACTGACATGATGTCAGAAATTCAG 1657
Oy 118 AlameGlnThrSerIleGlnGlnAlaSerLysSerMetGlySerThr---LeuGlySer 136
Db 1658 CAACCTGCAAGAGATTTTGCAGATCTACAGAAAGATTAGAGACTCTGAGACAGACAGA 1717
Oy 137 LeuGlnSerLysSerAlaAlaGlnMetLysGlyValGlnAlaValAlaVal---Ala 154
Db 1718 AAACAGCTACAGGTGCAACCTCCAGAGGCGACAGAACTGATGCTTAAACACACT 1777
Oy 155 AlaLeuSerGlyLysSerSerGly---SerAlaLysLeuGlnThrProGlnLeuProLys 173
Db 1778 GAGATTTCAGAGAACAGCTCTGACCTCAGCCAGAACTT-----AAA 1819
Oy 174 ProGlyValThrProArgSerGlyValIleGlnIleGlyLeuAlaLeuAlaLysAlaIle 193
Db 1820 GAACTTAGCAACAATACGAGAGGCTATGAAAGAGCTTAGTGTGACAGAGACAGATG 1879
Oy 194 GlnThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnThr--- 207
Db 1880 AAA---CTCGGCTCTTGTCTCAGCTGAAAGCATGATTAATTCACATTTCCACAGAGCTG 1936
Oy 208 -----AlaSerThrGln 211
Db 1937 AGGTCACGCAAGAGAAATTAATCTGCTAAAGCAGATCTGCAGAAATGCAATTAAAGAA 1996
Oy 212 AlaGlnAlaAspGlnThrAsnLysLeuGlnLysGlnAlaIleLysIleAspLys 231
Db 1997 AGTGAAGAAATTAAGAGAAAGTGAAGATGAGAGAACTGTGAGAGAGAGAGAAA 2056
Oy 232 GlnArg-----GlnGlnLysGlnGlnMetLysAlaAla----- 242
Db 2057 GGTACAGTGAATTAAGCCACCTGCTGAGAGATGACAGAAATGAATTAATTCATTTGCTCT 2116
Oy 242 ----- 242
Db 2117 GTTATTGAGAAATTAAGAGAAAGCATTTTGTGAGAAATACCAAGAACCCAA 2176
Oy 243 -----GlnGlnLysSerLys 247
Db 2177 GAAGAATCATGAAATTAAGACACACTTAAAGTCAGATGACACAGAGAACCCAGTGAT 2236
Oy 248 AspleuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIle 267
Db 2237 GAAGGTGAGACATGAAGAACCCATGAATGATGATGAATCAATTAACAGGTG 2296
Oy 268 ThrValIleSer----- 271
Db 2297 AGCGAGCTCTCACACTGTACAAAGAACCCAGCTGAGCTGAGAGATTACAGAAAGG 2356
Oy 272 -----IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAla 287
Db 2357 AAATCTTAGAGATGTCACAGCTGAATATATCCATTAAGCAGACATGAGAACTGATG 2416
Oy 288 AlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaAla 306
Db 2417 CAATTGACAAACGCTGCCAGGCTAAAGCAGAGATGCACTCTGAAATGAAGTCTCAG 2476
Oy 307 ---ThrThrValAlaAlaThrGlnIleThr---ValGlnAlaValAlaGlnAlaValLysGln 324
Db 2477 TATTCAAAAGTGTGTAATGAGTTGACCCAGCTCAAAACACTGTGATGACAAAAAAGG 2536
Oy 325 -----AlaValIleThrAlaValAlaGlnAlaIleThr 335

```

```

Db 2537 AACTCTGCTCTTATCACAGAACATTTTGCAGATGATTAACCAAGCTGCGG-----ACT 2587
Oy 336 AlaAlaIleLysAlaAlaValLys---SerGlyIleLysAlaPheIleLysThrLeuVal 354
Db 2588 GCAGCAAAAGAGATGAGAAAGAAAAATTAAGCAATCTTAAGAACACTTGCAGACAGAAAGAA 2647
Oy 355 LysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGln 374
Db 2648 GTGGAAGTACGAAGAGCTGAGAAACACTCTTACAGAAAGAAAGCTCTATGACTGATGCA 2707
Oy 375 MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTyr 394
Db 2708 ATGGTACCTCGGCTCT---TCTATGAAAAACTCCAGTCATCTTATGAGAGTGA--- 2758
Oy 395 ValThrValGlyValGlyValAlaValAlaAlaProAlaLeuGlyLysGlyIle----- 412
Db 2759 -----GTCAGTGTGTGTCATCGAAATTAAGGAATCTGTAAGAG 2800
Oy 413 -----MetGlnMetGlnLeuSerGlnLysGlnGlnAlaAlaGlnPheGlnLys 429
Db 2801 AAAGAGAGAGCTCATTCAGAGCTTGCAGATTAAGAGTGAAGTCTCAGAGTGAAGAAAGA 2860
Oy 430 Gln-----Val 431
Db 2861 GAAAGAGAAATTAATTCAGACTCTCTTGAATCCAAGAGCAAGAGTAAATGAACCTTCTG 2920
Oy 432 GlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnAla 451
Db 2921 CAAATATTCACGACAGACTCAGAGAAAGCTTCCAGAAATGAAGAAATTCGCGAGAGCTCT 2980
Oy 452 SerLysIleAlaSerLysGlnThrGlyGlySerAsnGlnMetThrGlnLysAlaThrLys 471
Db 2981 TCAAAACTGAGAGAAATTAAGATTAAGATTAATGATGATGAGTGAAGAGAGTCAACCAA 3040
Oy 472 Leu 472
Db 3041 TTG 3043

```

RESULT 17  
 LOCUS BC028951 2589 bp mRNA linear HTC 07-AUG-2002  
 DEFINITION Mus musculus, similar to E1A binding protein p400, clone  
 IMAGE:5364135, mRNA.  
 ACCESSION BC028951  
 VERSION BC028951.1 GI:22137754  
 KEYWORDS HTC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 2589)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amgdbcm.tmc.edu](mailto:amgdbcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 511 a 332 c 274 g 76 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 0.00408 Length: 1194  
Score: 136.00 Matches: 59  
Percent Similarity: 38.93% Conservative: 36  
Best Local Similarity: 24.18% Mismatches: 115  
Query Match: 5.74% Indels: 34  
DB: Gaps: 7

US-09-889-314-2 (1-496) x AL534303 (1-1194)

```

Oy 7 SerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThr 26
Db 427 TCTTCCACCCCAACCCACACCCCAAC----- 456
Oy 27 ProGlnGlyValProGlnGlnAsp-----LysLeuSerGlyAsnGlnThrLysGlnIle 44
Db 457 ---CAAGGCTTCACAGCAAGCAAGCGGGAAGCAAGCAAGCAAGCAAGCAAGCAAG 513
Oy 45 GlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
Db 514 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 573
Oy 65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
Db 574 AGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 633
Oy 85 AlaAlaGlyLysGlnSerSerGlnSerGln-----LysAlaGlyAlaAspThrGlyVal 102
Db 634 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 693
Oy 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 694 AGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 747
Oy 123 IleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla 142
Db 748 AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 807
Oy 143 AlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSerGlyLysSerSerGly 162
Db 808 GCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 834
Oy 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182
Db 835 GCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 894
Oy 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaIleThrLysSer 202
Db 895 AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 939
Oy 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeu 222
Db 940 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 999
Oy 223 GlnLysGlnAlaIleLys---IleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGln 241
Db 1000 AACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1059
Oy 242 AlaGlnGlnLys 245
Db 1060 GCAAGCAAGCAAG 1071

```

## RESULT 16

BC028681 3344 bp mRNA linear HTC 01-MAY-2002  
LOCUS Homo sapiens, retinoic acid induced 14, clone IMAGE:4822221, mRNA.  
DEFINITION BC028681  
ACCESSION BC028681.1 GI:20379528  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
DIRECT SUBMISSION  
Submitted (29-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshitsuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNU)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amadams@systemsbiology.org](mailto:amadams@systemsbiology.org)  
Anup Madan, Jessica Fahy, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

## REMARK

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNU at: <http://image.llnl.gov>  
Series: IRMA Plate: 33 Row: C Column: 23  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13470085  
This clone has the following problem: frame shifted.  
Location/Qualifiers  
1. 3344  
/organism="Homo sapiens"  
/db\_xref="LOCUSID:26064"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4822221"  
/issue\_type="Testis"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"

## FEATURES

## source

BASE COUNT 1170 a 691 c 782 g 701 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.0197 Length: 3344  
Score: 135.50 Matches: 113  
Percent Similarity: 37.10% Conservative: 110  
Best Local Similarity: 18.80% Mismatches: 207  
Query Match: 5.72% Indels: 171  
DB: Gaps: 23

US-09-889-314-2 (1-496) x BC028681 (1-3344)

```

Oy 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27
Db 1346 AGTACTACAGTGTCTAGTCTTATTCAGATTAAGTCTGTA----- 1387
Oy 28 GlnGlyValProGlnGlnAspLysLeuSer-----Gly 38
Db 1388 -----GCTGACCAACAGATCTTCTCTCTATTCAGCAAGCAAAAGTCTCCCTTACC 1441
Oy 39 AsnGlnThrLysGlnIleGlnGlnThrArgGlnGlnLysAsn---ThrGlnMetGlnSer 57
Db 1442 TTACACATTAAGAGGTTTACAGATTAATTCACGCCCAATCACCAGAGGCGGAGCA 1501
Oy 58 AspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlu 77

```





D	b	958	CCCTGGGAAAGGGGCTACCCACACACCCCTGGGAAGGACGGGGCTACCTCCACAC	101	17
O	y	121	rSerIleGluGluAlaSerIlySerMetGluSerThrIleGluIleSerLeu	141	11
D	b	1018	CAAGGACGGGAAGCCACAGAGGAGACTCAGAGACACACAGGAGAG--TCATCTGACAG	1074	11
O	y	141	rAlaAlaGluMetIySglValuAlGluAlaValAlaAlaIleuSerGlySerSe	161	11
D	b	1075	TGAGGAGAGACGCCACCTCCCAAGGCCCTGCTCAGGGGAAGGCCCTCAGAAAACTC	1134	11
O	y	161	r-----GlySerAlaIlyLeuGluThrProGluLeuProIlyValThrProAr	179	11
D	b	1135	TCAGGTGGAGACTGCTCCACACCCCTGCCAAGAGATGCCCCAGAAAGAGCTGCCCA--	1192	11
O	y	179	gSerGluValIleGluIleGlyLeuAlaLeuAlaIyAlaIleGluThrLeuGlyVal	199	11
D	b	1193	-GGCCCCCTGGGAGACAGCGGCTCAGTTCAGGACAGGCCAGCGGGGAAGCGGAGGA	1251	11
O	y	199	aThrIlySerAlaIleuSerIlyThrIaSerThrGluAlaGluAlaAsp-----	215	11
D	b	1252	GGACTCGCAGACGACGCGAGGAATCGACAGTAGAGAGGGCGCTGCTCAGCGCA	1311	11
O	y	215	-----	215	11
D	b	1312	GCCTTACAGGGAAGGCCCCCGACGCTCAGACGCCCTCGGCCCCAGAGACTCCCCAG	1371	11
O	y	216	-----GluThrAsnIlyLeuGlyLeuGluIySglAlaIleIyS	229	11
D	b	1372	GAAGGGGCTGCCCCACGACCTCTCTAGAAACAGGGCTCGACGCCGCCACGTCAGT	1431	11
O	y	229	eAspIySglIuArgGluGluIuTrGluIleuMetIySAlaAlaGluGluIySerIyAsp	249	11
D	b	1432	GGGGAACACAGAGAGACACCAAGAACGACGACGAGAGTCAACAGTACAGAGAGAC	1491	11
O	y	249	uGluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVa	269	11
D	b	1492	ACTGGACGCCATGAATCCACTCAGGTGAAGCCTTGGGGAAAGCCCCAGGTAAACC	1551	11
O	y	269	IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGl	289	11
D	b	1552	TGCTCT-----ACATTGGCATGGGGCTTGGGGAAGGGCGCG	1593	11
O	y	289	y-----AlaAlaValoIyAlaAlaAlaAlaAlaGlyAlaAlaGly-----	302	11
D	b	1594	CCGAGTCCACCGGGGAAGGTGGGCTGTCAACCCCTCAGCCAGTGGGGAAGTGGGA	1653	11
O	y	302	-----	302	11
D	b	1654	GGAGACTCTCAGAGAGCAGTAGTAGAGATCATCAGACAGCACTGATGAGAGTCCAC	1713	11
O	y	303	-AlaAlaAla-----	305	11
D	b	1714	AGCTGTGGCCCCCGCTCAGAAAAAGTCCTTGGGAACATCTCCAGGCCAAMCCACCTC	1773	11
O	y	306	-----AlaThrThrValAlaIleThrGluIleThrValGl	316	11
D	b	1774	CAGTCTGCCAAGGGGCCCTCAGAAAGCGAGGCTGTAGCCCTCAGCTCAAGTCAAGCTGA	1833	11
O	y	316	nAlaValAl-----GluAlaValIySglAlaValIleThrAlaValAlaArgAl	333	11
D	b	1834	AAAGCCCATGGACACTCTGGAGAGACGCGAGGAGTATCGACAGTGGACAGTGAAGA	1893	11
O	y	333	alleThrAlaAlaIleIySAlaAla--ValIySerGlyIleIyS-----	347	11
D	b	1894	GGCAGCAGCAGCGCATGTCACTCAGCTCAGGCAAAACAGCTGTGAATAATCTCCACACCA	1953	11
O	y	348	-----AlaPheIleIySThrLe-----	353	11
D	b	1954	GGCTGCCCCAAGAAACCAATACACATGCACTGTGCCAAGTGTGCCCTGTGCGAGTGG	2013	11
O	y	353	uValIySAlaIleAlaIySAla-----IleSerIySgIlyIleSerIyValPh	369	11
D	b	2014	CACCCAAAGCCCCGGAAACAGAGAACTGCGACTTTCACAGAGGCTCATTCACAGCTGT	2073	11

Oy	369	ealalsylslyhrnglnmetllealalsasnpheroysleuserlysvallieser	369
Db	2074	GGCTGGGGGACCCACGACGACGAGATTCTTCAAGCAGTGAAGATCATGATATGCA	2133
Oy	369	rleuthserlystrpvaltrnvalglvalcglyvalvalvalalalaprovalencl	409
Db	2134	GGAAGAGAAAGACAGGTCTTGCAAGTAAACCTGGGA-----CAGGCAAGCTGTGGG	2184
Oy	409	yllyscglyllemetclnmetglinleuser--glumetclnlnasnvalalaglnphecl	428
Db	2185	GAAGAGGCTCCACGAGGAAGACAGCTCAGCTCTCAAGGGGTCTTGCGGCAAGGCAC	2244
Oy	428	nlyslval1-----glytysleuaglnalalalalalspmetllesermetphethcl	446
Db	2245	TGCTCCACTACTCTCCCTGGGAGAGCGGGCTCAAGTCACCCAGG-----AA	2292
Oy	446	nphetpncinglnalalsaserlylealalseryglnthrlygluserasnglumetth	466
Db	2293	AGCTGAAGAAGACGAGAGACTGTGAGACGAGTGAAGAGAAATCAGACAGTGAAGAACGC	2352
Oy	466	rglnlysalathrlystleuglyalaglnlleleulyslatryalalalalleserglyal	486
Db	2353	TGCATCTCCACACAGAGTGTAAGAACTCTGATTAAGAAACCCAGGCCAAAGCCACCGAC	2412
Oy	486	alleaglaalyala	490
Db	2413	TGCGCGCAGACGA	2425
RESULT 13			
LOCUS	A2624776	746 bp	DNA linear
DEFINITION	1M0463K1R Mouse 10kb plasmid UUGC1M library Mus musculus genomic		
ACCESSION	A2624776		
VERSION	A2624776.1	GI:11747062	
KEYWORDS	GSS.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 746) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.		
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts		
JOURNAL	unpublished (2000)		
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert length: 10000 Std Error: 0.00 Plate: 0463 row: K column: 14 Seq primer: CACACAGAAAACAGCTATGACC Class: plasmid ends High quality sequence stop: 746.		
FEATURES	Location/Qualifiers		
SOURCE	1..746		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="UUGC1M0463K14"		
	/clone_id="Mouse 10kb plasmid UUGC1M library"		
	/sex="Male"		
	/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"		
	/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson		

```

QY 250 GIUGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVal 269
      |||||
Db 842 -----GTTATATATCCCGGAAAGGTTGGCGT----- 868
QY 270 IleSerIleValAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGly 289
      |||||
Db 869 -----GTCGCGGATACGGGTGATGTGCGC----- 892
QY 290 AlaAlaValAlaIleAlaAlaGlyGlyAlaGlyAlaAlaAlaIleThrVal 309
      |||||
Db 893 -----AAGGTTGCTGCTGCTGCCCAAGCAGCTGCT-----GCCCGTGCATTT 937
QY 310 AlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVal-----IleThr 328
      |||||
Db 938 GTGACCGAGATTCAGCCCATCTGTGCTCCAGGCTCTGTGAGAGGGTCTTCAGGCTCTT 997
QY 329 AlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAla 348
      |||||
Db 998 CCTTGAGAGAGCTGTCTGTGAGCTGACATCTGTGACCACTGCGCAACAGAT 1057
QY 349 PheIleLysThr-----IleuValLysAlaIleAlaLysAlaIleSerLysGlyIleSer 366
      |||||
Db 1058 ATCATCATGCTTGACCATGAGAGATGAGACATGCTCATTTCTGCAACATTTGGC 1117
QY 367 LysValAlaPheAlaLysGlyThrGlnMetIleAlaLysAsnProLysLysSerLysVal 386
      |||||
Db 1118 CACTTGCACATGAAATGATGATGCTGCGCTTGAGACCTTACCTGCGCAAGCGCATC 1177
QY 387 -----IleSerSerLeuThrSerLysTyrVal-----ThrValGlyValAlaValAl 403
      |||||
Db 1178 ACATCATAGCCCAAGCATGACCGCTGGGTGTCCCGCAGACCAACACTGGCATCATTTGTC 1237
QY 404 AlaAlaPro-----AlaLeuGlyLysGlyIleMetGlnMet 415
      |||||
Db 1238 CTTCGCTGAGGCTGCGCTGATGAACTTGGTGTCTACTGCTGACCTTGTTCATG 1297
QY 416 GlnLeuSerGlnMetGlnGlnAsnValAlaGlnPhe-----GlnLysGlnValAl 431
      |||||
Db 1298 TCCTGCTCATTTCACTACCAAGGTCATTTGCCCACTTGAATGAGAGGAGAGAGCTCT 1337
QY 432 GlyLysLeuGlnAlaIleAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnAla 451
      |||||
Db 1358 GGCAGATGATGAGAAAGAGGTATGTGCTCCCAAGCACTTGAT----- 1402
QY 452 SerLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLys 471
      |||||
Db 1403 GAGAAAGCTTCTGCTCTCCACTTGGC-----AAG 1432
QY 472 LeuGlyAlaGlnIleLeuLysAlaThrAlaAla-----IleSerGlyAlaIle 487
      |||||
Db 1433 CTTCGCTCCAGCTGACCAAGCTCACCAGCTCTAGGCCGACTACATCAGCGTCCGATC 1492
QY 488 AlaGlyAlaIleLysIleThrAsnAspPhe 496
      |||||
Db 1493 GAGGCTCCCTAACAAGCTGCCCACTAC 1519

RESULT 12
LOCUS BC016144 4272 bp mRNA linear HTC 24-OCT-2001
DEFINITION Homo sapiens, clone IMAGE:3921754, mRNA.
ACCESSION BC016144
VERSION BC016144.1 GI:16359374
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

```

REMARK
COMMENT
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
InfoBioscience.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chen, Readman Chiu, Chris Fjell, Erin Garland, Ran Glin,
Leticia Hsiao, Martin Krzywinski, Reto Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candace McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Pihlhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRM Plate: 15 Row: n Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507410
This clone has the following problem: frame shifted.
Location/Qualifiers
FEATURES
source
1..4272
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3921754"
/tissue_type="SKIN, melanotic melanoma."
/clone_lib="NIH-MGC_72"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
BASE COUNT 1156 a 1221 c 1352 g 543 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00987 Length: 4272
Score: 140.00 Matches: 118
Percent Similarity: 33.12% Conservative: 89
Best Local Similarity: 18.88% Mismatches: 257
Query Match: 5.91% Indels: 162
DB: 11 Gaps: 19

US-09-889-314-2 (1-496) x BC016144 (1-4272)
QY 7 SerSerSerSerGlyProAspAsnGlnLysAsn-IleMetSerGlnValLeuThrSerTh 26
      |||||
Db 614 GCCGACAGCTCCACGAGACACCTCCAGCTCCAGTATGAGACAGAGGAGGAGAA 673
QY 26 rPrGlnGlyValProGlnGlnAspLysLysSerGlyAsnGlnLysGln----- 43
      |||||
Db 674 CCGTCAGTA-AAACCAAGCCAGGTCAAGCCCTATCAGTTTCTACTAGAGAGTCTCCAGC 732
QY 43 ----- 43
      |||||
Db 733 AAGAAAGGCGGCCCGCCCTGGAAGAGTGGGAGTGTGACACCCAGTCAAAGAGAG 792
QY 44 -----IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAl 62
      |||||
Db 793 GGCCCTGCCCCCAGCAAGAGGCGCAAGAACCAAGAGAGAGTCAAGAGTACTAGAGA 852
QY 62 aglyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaProGlnG 82
      |||||
Db 853 GGGATCTGAAGTAGAGGAGGCCCTGCGAGGACGAGAGCCAGGTAAGGCTCTGCA 912
QY 82 nGlyValAlaIleLysGlyLysSerGlnLysAlaGlyAlaIleAspThrGlyVal 102
      |||||
Db 913 AAAAATTCAGAGTCAAGAGTCTGCTCAGCCCTGCCAAG-----GGGAC 957
QY 102 lSerGlyAlaAlaIleThrAla-----SerAsnThrAlaThrLysIleAlaMetGlnTh 121
      |||||

```

Accession	Gene	Protein	Length	Source
Oy 127	serlysermetglusertphreugluserleugluser	139		
Db 1902	GCAGAGCCCCAGTCACAGGACCAACAACATGAGGCGGCTTAGGCCACAGG	1961		
Oy 140	leuseralaalaglmetlysgluval	156		
Db 1962	ATGGAAGCCACAGAGTCAATGTCACAGCTGTGGAAACCAAGGCTGAGAGCGGTC	2021		
Oy 157	serglysserserglyseralalyseugluthrprogluleuprolypsproglyval	176		
Db 2022	CCGGGATCTCTGCTGCGC-----CCCATCTCAATCTGCTGCTGC	2060		
Oy 177	thrprougsergluvalilleglulidelyleualaleualalyalaillleglthrleu	196		
Db 2061	GCAGAGGCTCTGGAAAGACCTTCACATA-----GAGCTGGAGACCAGATCCGTGCTTG	2114		
Oy 197	glyluatalfhrlysserseralaleuserantfyralsertfhrglinalaglinalaspcln	216		
Db 2115	GAGGAGGCTCTCCGCGAC-----CGGAGCGGAGAGCGACGTCCGAGCTG	2159		
Oy 217	thrsnlylsleugly--leuglulysglinalalylealyleasplysgluarglugu	235		
Db 2160	GAGGCGGCTGTGGAGATGTCGAGGCCCGGAGCGCAGAGCGCGGCTGTGAGAGCGT	2219		
Oy 236	tyrtinglumetlys-----	240		
Db 2220	GTCCGCGAGCGCCAGGAGCGAGCGGCGGCGGTGGCGGTGACACACAGCTG	2279		
Oy 241	-----AlaAlaaglInglLylsersylsAspleugluguYlthrmetaSptfhrValasnthr	258		
Db 2280	CGGCGGCTGTGGAGATGTCGAGGCCCGGAGGAGACCTCCGACCGGAGCTCCGCTGCGGAG	2339		
Oy 259	valmetilealavalservalalalierthralilleserilevalalalalierpether	278		
Db 2340	CTGAGGACGCTCGCTGCTGATGATGAGGCTCGGCGAC--CGGCTGTGCGGAGGA	2398		
Oy 279	CysglYAlaGlYleuAlaGlYleuAlaAlaGlYAlaAlaValaGlYAlaA-----	295		
Db 2399	GGAGCGCGGCGGCGCTCGCGCGAGCTGCGCCAGCGGAGGAGCGCGGCTGTGAGCAGAG	2458		
Oy 296	---AlaAlaGlYAlaAlaGlYAlaAla 304			
Db 2459	CCGGAGCTGGAGGTTCTGCGGAGACGCT	2488		
RESULT 11				
LOCUS	AY103583	1845 bp	linear	HTC 25-MAY-2002
DEFINITION	Zea mays PCO113569 mRNA sequence.			
ACCESSION	AY103583			
VERSION	AY103583.1			
KEYWORDS	HTC.			
SOURCE	Zea mays.			
ORGANISM	Zea mays.			
REFERENCE	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACC clade, Panicoideae, Andropogoneae, Zea.			
REFERENCE	1 (bases 1 to 1845)			
AUTHORS	Hanley,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.			
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	2 (bases 1 to 1845)			
AUTHORS	Coe,E.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			
FEATURES	Location/Qualifiers			
source	1..1845			
	/organism="Zea mays"			
	/db_xref="MaizeDB:63617"			
	/db_xref="taxon:4577"			
	/clone="PCO113569"			

[illegible]



```

Db      208  TCCAAATGCATCTCCACACTTCTGGTGTTCATCTGCTTCTGATCTAGTATCCGAT---- 264
Oy      96  AAGAGTAAAspThrglyValserGlyAlaAlaThrAlaSerAsnThrAlaThr 115
Db      265  GCTTCCTGACCTCTGCTCTGCTCTGATGCTAGTGAATCCAAATGCATCTGCACCTCTGCT 324
Oy      116  LysTleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGlnSerThrLeuGlu 135
Db      325  -----GTTTCACTGCTCTCTCTGATGCTAGTGAATCCAAATGCATCTGCACCTCTCT 372
Oy      136  SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaValAlaAla 155
Db      373  GGTGTTTCACTCTGCTCTGCTCTGATGCTAGTGAATCCAAATGCATCTGCTGCTCTCT 432
Oy      156  LeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGly 175
Db      433  GATGCTAGTGAATCCAAATGCATCTGCC-----ACTTCTGCT 468
Oy      176  ValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195
Db      469  GTTTCATCTGCTCTGCTCTGCTAGTGAATCCAAATGCATCTGCACCTCTGCTGTTTCTCT 528
Oy      196  LeuGlyGluAlaThrLysSer-----AlaLeuSerAsnThrAlaSerThrGlnAlaGln 213
Db      529  GCTTCGATGCTAGTGAATCCAAATGCATCTGCCACTTCTGCTGCTCTGCTCTGAT 588
Oy      214  AlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArg 233
Db      589  GCTAGTGAATCCAAATGCATCTGCCACTTCTGCTGCTCTGCTCTGCTAGTGAATGCT 648
Oy      234  GluGluThrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMet 253
Db      649  TCCGATGCTCTGCTACCTCTGCTGCTCTGATGCTAGTGAATCCAAATGCATCTGCCACT 708
Oy      254  AspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 273
Db      709  TCTGCTGTTTCACTCTGCTCTGCTGCT-----AGTGAATCC 744
Oy      274  AlaAlaIlePheThrGlyGlyAlaGlyLeuAlaGlyLeuAla-----AlaGly 289
Db      745  GATGCTTCTGCCACTTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
Oy      290  AlaAlaValAlaGlyAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaThrThr 308
Db      805  GCCACTTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
Oy      309  ValAlaThr 311
Db      865  GGTGCTACA 873

RESULT 10
LOCUS      BC028236          2524 bp      mRNA      linear      HTC 01-MAY-2002
DEFINITION Homo sapiens, similar to RIKEN cDNA 1700020L11 gene, clone
IMAGE:5245129, mRNA.
ACCESSION  BC028236
VERSION    BC028236.1 GI:20380168
KEYWORDS   HTC.
SOURCE      Homo sapiens.
ORGANISM   Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2524)
            Strausberg, R.
            Direct Submission
            Submitted (08-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK     NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT    Contact: MGC help desk
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.

```

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nsl.llnl.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Latic, P., Legaspi, K., Maduro, Q. L.,
Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J.,
Pearson, R., Stantripop, S., Thomas, P. J., Touchman, J.W., Tsurgueon, C.,
Vogt, J. L., Walker, M.A., Wetherby, K.D., Wiggin, L., Young, A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 62 Row: n Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.

FEATURES
    source
        1..2524
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5245129"
        /tissue_type="brain, fetal, whole pooled"
        /clone_11b="N1H-MGC_121"
        /lab_host="DH10B"
        /note="Vector: PCMV-SF0R6"

BASE COUNT      588 a      714 c      897 g      325 t
ORIGIN
Alignment Scores:
Pred. No.:      0.00296      Length:      2524
Score:          142.00      Matches:      85
Percent Similarity: 38.57%      Conservative: 50
Best Local Similarity: 24.29%      Mismatches: 134
Query Match:      5.99%      Indels:      82
DB:              11      Gaps:      13

US-09-889-314-2 (1-496) x BC028236 (1-2524)
Oy      2  ThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21
Db      1542 ACCACATGAGCTAAATGCTCAGTGGCTCCGAAACCAAGTTAAAGGAGCCGAGACC 1601
Oy      22  ValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThr 41
Db      1602 ATA-----GATGAGAGAGGCTCGAGAGATGAAC 1631
Oy      42  LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGlnSerAspAlaThrIle 61
Db      1632 ATGGAACCCAG-----ACTATGGAACCTGAGGCGCCAGCGCA 1667
Oy      62  AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGln 81
Db      1668 GCCAAGGCCACGGAGCTGAGGCCACAGGCCAAGGTC---ACAGAAACAAACCAACCCACA 1724
Oy      82  GlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGly 101
Db      1725 GGGGCTGAGCTCAGAGAAATGAGAGACACAGAA-----GAAGAACCAACATGAACTAAG 1781
Oy      102  ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrIle----- 117
Db      1782 CCCACGAGACTCAGGCCAGACACAGACAGACACGAGAGGAGGCCATGGGGGTGGAG 1841
Oy      118  AlaMetGlnThrSerIleGlnGluAla----- 126
Db      1842 GCCACAAACCAAAAGACAGAGGAAACGAAATGACAGGCTACGAGATGGGTGCTGGGCA 1901

```





5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

# FEATURES

source

1. 1094  
/organism="Pichia angusta"  
/strain="CBS 4732"  
/db\_xref="taxon:4905"  
/clone="BB0AA002F11"  
/clone\_1lb="BB0AA"  
/note="end : r3"  
BASE COUNT 320 a 245 c 321 g 207 t 1 others  
ORIGIN

## Alignment Scores:

Prod. No.: 0.000255 Length: 1094  
Score: 147.50 Matches: 75  
Percent Similarity: 39.19% Conservative: 70  
Best Local Similarity: 20.27% Mismatches: 174  
Query Match: 6.23% Indels: 51  
Gaps: 9

US-09-889-314-2 (1-496) x CNS076CM (1-1094)

QY 8 SerSerSerGlyProaspanglnlysaSnllleMetSerGlnValleuThrSerThrPro 27  
DB 1008 AGCCAAAGCTCAACTTCCAAACAGGCTTCAGGTGATCCAGCTCTACTGCTGCTTCA 949  
QY 28 GlnGlyValProGlnGlnAspLysLeu-----SerGlyAsn 39  
DB 948 AGTAATGCTCTGACCTGATACCATTCACAGATCTCCGTCAGTAGTGGATGGGCTCG 889  
QY 40 GlnThrLysGlnleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspala 59  
DB 888 TCCAGCGGTTCTAGCTTCGACAGGTTATGATCTAGCACTGCTATCCGATTCCGA 829  
QY 60 ThrIleAlaGlyLaserGlyLysAspLysThrSerSerThrThrLysThrGlnThrAla 79  
DB 828 ACTTTCAGCTGCTCT-----ACCTCAAGCTCAGGTTTATGCTCAAGGCC 784  
QY 80 ProGlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAsp 99  
DB 783 GCCACCTCGATTTATGATGCTTACTTCAAGCGGTCAGGTCACGCTTCCAGC 724  
QY 100 ThrGlyAlaSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMet 119  
DB 723 TCTGCCACTCCTCGATCAACGCTTTCAGCTGCTGCTCCGCGCTACT 664  
QY 120 GlnThrSerIleGlnGlnAlaSerLys-----SerMetGlnSerThrLeu 134  
DB 663 TCTGCCCTCAAGTTATGATCAACACAGCTAGCTCTGTTAGCTCGCGCTGCTATA 604  
QY 135 GlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAla 154  
DB 603 TCAGGTTTGAAGCTCTGCTCAAGTTATGATCAAGGACTACTGCTAGTTTAAAGCTCT 544  
QY 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro 174  
DB 543 GCGGCTCTGCTTCCAGTTATGCTCAAGC----- 514  
QY 175 GlyValThrProArgSerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln 194  
DB 513 -----TCTACTGCTGCTGCTTGAAGCTCTGCTCAAGTTATGCGTTAAGC 469  
QY 195 ThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAla 214  
DB 468 ACCGAGGCTGCTGCT-----TCAAGTTATGATCAAGACCGCAAGCTCT 424  
QY 215 AspGlnThrAsnLysLeuGlyLeuGlnLys-----GlnAlaIleLysIleAspLysGlnArg 233  
DB 423 -----GGTTTGAAGCTCTGCGCTGCTCAATATAGGCTCAACTCTGCC 382  
QY 234 GlnGlnLysThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGly---Thr 252

DB 381 TCAAGCTATGCGCTCAAGCACTGCTGCTCTACTTTAAGCTTCGCGCTGCTGCTCACT 322  
QY 253 MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIle 272  
DB 321 TATGCTCAAGTTCTGCTCTGCTGATCAAGCTATGCTCAAGCTCTGCTGCTCAAGC 262  
QY 273 ValAlaAlaIleThrLysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaVal 292  
DB 261 TCCAGCTCTGCTCAAGTTATGATCAAGCACTGCTGCTGCTGCTCAAGCTCTGCTCA 202  
QY 293 GlyAlaAlaAlaAlaGlyAlaGlyAlaAlaAlaAlaAlaAlaAlaThrValAlaThrGln 312  
DB 201 AGTTATGCTCAAGCTCTGCTCTGCTCAAGCTCAAGCTCTGCTCAAGCTATGCTCAAGC 142  
QY 313 IleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGln 332  
DB 141 TCTACTGCTCTGCTGCTGCTGCTCAAGTTATGCTCAAGCAAGCTCAAGCAAGCAAGC 91  
QY 333 AlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaIleThrLysThr 352  
DB 90 -----GGCTCAGGTTTGAAGCTCTGCTCAAGTTATGATCAAGCTCTACTGCTGCT 37  
QY 353 LeuValLysAlaIleAlaLysAlaAlaIleSer 362  
DB 36 TTTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7  
RESULT 8  
BC016304/c 2295 bp mRNA linear HTC 05-NOV-2001  
LOCUS Homo sapiens, similar to dentin staphosphoprotein, clone  
DEFINITION IMAGE:3844335, mRNA.  
ACCESSION BC016304  
VERSION BC016304.1 GI:16740886  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Strausberg, R.  
AUTHORS Direct Submission  
TITLE Submitted (29-OCT-2001) National Institutes of Health, Mammalian  
JOURNAL Gene Collection (MGC), Cancer Genomics office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA.  
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov  
COMMENT Contact: MGC help desk  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
contact: amadan@systemsbiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
Series: IRAL Plate: 25 Row: b Column: 20  
This clone has the following problem: frame shifted.  
location/Qualifiers  
1. 2295  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3844335"  
/tissue="IMAG:3844335", choriocarcinoma"  
/clone\_1lb="NIH-MGC-21"  
/lab host="DH10B-R"  
/note="Vector: pOTB7"  
BASE COUNT 558 a 732 c 560 g 445 t  
ORIGIN

Eukaryota: Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;  
 Eimeria.  
 1 (bases 1 to 619)  
 Ng,S.T., Jangli,M.S., Shirley,M.W., Tomley,F.M. and Wan,K.L.  
 Comparative EST analyses provide insights into gene expression in  
 two asexual developmental stages of *Eimeria tenella*  
 Unpublished (2001)  
 JOURNAL  
 COMMENT  
 Contact: Wan K.L.  
 Centre for Gene Analysis and Technology  
 Universiti Kebangsaan Malaysia  
 43600 UKM Bangi, Selangor DE, Malaysia  
 Tel: 6 03 8292997  
 Fax: 6 03 8293249  
 Email: klawan@krisc.cc.ukm.my  
 PCR Primers  
 FORWARD: T3  
 BACKWARD: T7  
 Seq primer: SK.

FEATURES  
 source  
 Location/Qualifiers  
 1..619  
 /organism="Eimeria tenella"  
 /strain="Houghton"  
 /db\_xref="taxon:5802"  
 /clone="etsHd17"  
 /clone\_id="ETC11"  
 /dev\_stage="Sporozoite"  
 /lab\_host="X11-Blue MRF"  
 /note="Vector: Lambda ZAPIT; Site\_1: EcoRI; Site\_2: XhoI;  
 Sporozoites were excysted in vitro from E. tenella H  
 oocysts, purified by column chromatography and mRNA  
 extracted using a FASTTRACK kit (Invitrogen). cDNA was  
 synthesised and a uni-ZAP XR library was constructed  
 using cDNA synthesis kit, ZAP-cDNA synthesis kit and  
 ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The  
 library was amplified once through E. coli X11-Blue  
 MRF."

BASE COUNT 222 a 163 c 180 g 54 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.000114 Length: 619  
 Score: 147.50 Matches: 69  
 Percent Similarity: 36.59% Conservative: 21  
 Best Local Similarity: 28.05% Mismatches: 106  
 Query Match: 6.23% Indels: 51  
 DB: 13 Gaps: 5

US-09-889-314-2 (1-496) x BG929588 (1-619)

OY 252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271  
 DB 111 ::: 111 111 111  
 DB 12 ACCAGGGAAGCAGCAAGCTGTACAGCAGCAGCAAGGAGAGCAGAC 71  
 OY 272 TleValAlaIleIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIle 291  
 DB 111 ::: 111 111 111 111 111 111  
 DB 72 GATGACGACAGACAGCAACACCAAGCTGCAAGAGCAGCAGCACTGACG 131  
 OY 292 ValGlyAlaIleAlaIleAlaIleGlyAlaIleGlyAlaIle-----AlaAlaIle 306  
 DB 111 111 111 111 111 111 111 111 111 111  
 DB 132 AACGCTGACAGCAAGAACTGCAGCAAGAACTGCAGCAAGAACTGTACCAAAAGCTGCAGCA 191  
 OY 307 ThrThrValAlaIleThrGlnIleThrValGlnAlaValAlaGlnAlaValGlnAlaVal 326  
 DB 111 111 111 111 111 111 111 111 111 111  
 DB 192 AGGACTACGACCAAAAGCTGCAGCAAGAGCTGCAGCAAGAGCTGCAGCA 251  
 OY 327 IleThrAlaValAlaGlnAlaIleThrAlaAlaIleLysAlaAlaVal-LysSerGlyTil 346  
 DB 111 111 111 111 111 111 111 111 111 111  
 DB 252 AGGACTGACAGCAAGAACTGCAGCAAGAGCTGCAGCAAGAGCTGCAGCA 311  
 OY 346 elysAlaPheIleIleThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSe 366  
 DB 111 111 111 111 111 111 111 111 111 111  
 DB 312 AAAAGCTGCAGCAAGAGCTGCAGCAAGAACTGCAGCAAGAGCTGCAGC 371

OY 366 rlyValPheAlaLysGlyThrGlnMetIleAlaLysAsnDheProLysLeuSerLysVa 386  
 DB 372 AAAAGCTGCAGCAAGAACTGCAGCAAGAACTGCAGCAAGAACTGCAGCA 411  
 OY 386 lIleSerSerLeuThrSerLysTTPValThrValGlyAlaGlyValAlaIleA 406  
 DB 412 -----GCTGCAGCGGCAAGCC 428  
 OY 406 CalareuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGln--GlnAsnValAl 425  
 DB 429 AGCAACCGGAGAGAGATGGTGAATTTCTGCTGCAGAGCAGCAGCAACCTGCAGCA 488  
 OY 425 agLpHeGlnLysGlnValGlyLysLeuGlnAlaAlaIleAspMetIleSerMetPheTh 445  
 DB 489 GCAGTGGCAGCAGGAG----- 504  
 OY 445 rGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMe 465  
 DB 505 -----CAGCAACTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 554  
 OY 465 thrGlnLysAlaThrThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerG 485  
 DB 555 GCTGCAGCA-GCA-----GCAGCAGTATCGGAGAGCCGCTGCAGCAGCTGCAGC 601  
 OY 485 yAlaIleAlaGlyAla 490  
 DB 602 AGCAGCAGCAGCAGCA 617

RESULT 7  
 CDS076CM/C 1094 bp DNA linear GSS 07-JUL-2001  
 LOCUS T3 end of clone BB04A002F11 of library BB04A from strain CBS 4732  
 DEFINITION of *Pichia angusta*, genomic survey sequence.  
 ACCESSION AL431228  
 VERSION AL431228.1 GI:12214640  
 KEYWORDS GSS.  
 SOURCE *Pichia angusta*.  
 ORGANISM *Pichia angusta*.  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; *Pichia*.

REFERENCE  
 AUTHORS Souciet,J.L., Aigle,M., Attiguenave,F., Blandin,G.,  
 Bolotin-Fukuhara,M., Bon,E., Brodier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durieux,P., Leplingle,A., Llorente,B.,  
 Malpertuy,A., Neugeglise,C., Ozier-Kalogeropoulos,O., Potier,S.,  
 Saurin,W., Tekala,F., Toffano-Nicche,C., Wesolowski-Louvel,M.,  
 Wincker,P. and Weissenbach,J.  
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
 yeast species for molecular evolution studies  
 FEMS Lett. 487 (1), 3-12 (2000)  
 TITLE 2 (bases 1 to 1094)  
 JOURNAL Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Attiguenave,F.  
 and Dujon,B.  
 Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia*  
*angusta*  
 FEMS Lett. 487 (1), 76-81 (2000)  
 MEDLINE 20584723  
 PUBMED 11152888  
 REFERENCE 3 (bases 1 to 1094)  
 TITLE Direct Submission  
 JOURNAL Genoscope.  
 Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
 sequele@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT This GSS is part of a random genomic sequencing program of thirteen  
 yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to

```

OY 273 ValAlaAlaIlePheThrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaVal 292
    |||||
Db 238 GTAGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 297
OY 293 GYAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 312
    |||||
Db 298 GTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
OY 313 IleThrValGlnAlaValAlaGlnAlaValAlaValAlaValAlaValAlaValAlaVal 332
    |||||
Db 355 ---GTTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
OY 333 AlaIleThrAlaAlaIleAlaAlaAlaValAlaValAlaValAlaValAlaValAlaVal 351
    |||||
Db 412 GTTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
OY 352 ThrLeuValAlaAlaIleAlaAlaAlaIleSerLys---GlyIleSerLysValAlaPheAla 370
    |||||
Db 472 GCTGTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
OY 371 LysGlyThrGlnMetIle-----AlaLysAsnProLysLeuSerLysValIleSer 388
    |||||
Db 532 GCTGCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
OY 389 SerLeuThrSerLysTrpVal---ThrValGlyValGlyValAlaValAlaAlaAlaAla 407
    |||||
Db 592 GCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651

RESULT 5
BM588321 711 bp mRNA linear EST 25-FEB-2002
LOCUS 1700687322518 A.Gam.ad.cdna.blood1 Anopheles gambiae cdna clone
DEFINITION BM588321
VERSION 19600449697628 5', mRNA sequence.
KEYWORDS EST.
ORGANISM African malaria mosquito.
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 711)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
TITLE 'R', Collins,F.H., Venter,J.C., and Hoffman,S.L.
JOURNAL Celera Anopheles gambiae EST project
COMMENT Unpublished (2002)
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004AAR row: F column: 14
Seq primer: M13 Reverse.
FEATURES
Source location/qualifiers
1..711
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449697628"
/clone_lib="A.Gam.ad.cdna.blood1"
/dev_stage="Adult"
/lab_host="DH10b"
/note="Vector: pSport1; site_1: salI; site_2: NotI. whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cdna inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mt4.org)"
BASE COUNT 305 a 234 c 135 g 37 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.00123 Length: 711
Score: 148.00 Matches: 64
Percent Similarity: 38.55% Conservative: 32
Best Local Similarity: 25.70% Mismatches: 134
Query Match: 6.25% Indels: 20
DB: 13 Gaps: 4

US-09-889-314-2 (1-496) x BM588321 (1-711)
OY 70 ThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyValAlaAlaGlyLysGlu 89
    |||||
Db 4 ACCGCTCCGGCAACACACACACACACACACACACACACACACACACACACACACACACACAC 57
OY 90 SerSerGlnSerGlnLysAlaGlyAlaSerThrGlyValSerGlyAlaAlaAlaThrThr 109
    |||||
Db 58 -----ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 102
OY 110 AlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnLysSerLysSer 129
    |||||
Db 103 GCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 162
OY 130 MetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlu 149
    |||||
Db 163 AGAAATGCACTTGCCTGCTCAACAAAACAGCAGCATCAGTGCA----- 204
OY 150 AlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 169
    |||||
Db 205 -----GCATTCACACATGCTCCAGGAAATTCACACATGCTCCAGGAAATTCACACATGCTCCAGGAA 258
OY 170 GluLeuProLysProGlyValThrProLysSerGluValIleGlnIleGlnLeuAlaLeu 189
    |||||
Db 259 TCA-----AATGCACACATTCACACATTCACACATTCACACATTCACACATTCACACATTCACAC 312
OY 190 AlaLysAlaIleGlnThrLysGlnGlyAlaThrLysSerAlaLeuSerAsnTrpAlaSer 209
    |||||
Db 313 ACACGACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 372
OY 210 Thr-GlnAlaGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 229
    |||||
Db 373 ACACGACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 432
OY 229 eAspLysGluArgGlnGluLysThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLe 249
    |||||
Db 433 ACAACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 490
OY 249 uGlnGlyThrMetLysPheValAsnThrValMetIleAlaValSerValAlaIleThrVa 269
    |||||
Db 491 -CAGCA-ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 548
OY 269 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGln 289
    |||||
Db 549 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 608
OY 289 yAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAla 309
    |||||
Db 609 AACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 668
OY 309 lAlaThrGlnIleThrValGlnAla 317
    |||||
Db 669 AACATCTGCACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 693

RESULT 6
LOCUS BG929588 619 bp mRNA linear EST 31-DEC-2001
DEFINITION eUSHEST0008 Eth11 Eimeria tenella cdna clone etsha017 5', mRNA
sequence.
ACCESSION BG929588
VERSION BG929588.1 GI:18002978
KEYWORDS EST.
SOURCE Eimeria tenella.
ORGANISM Eimeria tenella

```

FEATURES

Location/Qualifiers

1. 935

/organism="Tetraodon nigroviridis"

/db\_xref="taxon:99883"

/clone="208D22"

/clone\_lib="G"

/note="Genoscope sequence ID : C0AG208DB11SP1-end : PUC-Or1"

BASE COUNT 136 a 251 c 273 g 271 t 4 others

ORIGIN

Alignment Scores:

Pred. No.: 5.12e-05 Length: 935

Score: 153.50 Matches: 65

Percent Similarity: 38.58% Conservative: 33

Best Local Similarity: 25.59% Mismatches: 131

Query Match: 6.48% Indels: 26

DB: 17 Gaps: 4

US-09-889-314-2 (1-496) x CNS033D4 (1-935)

OY 147 GluValGluAlaValValValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeu 166

DB 793 GAAGCTGAAGCTCTGTGTCTCTCTCTGCTGCTGGGTGACGATGCGCGGACACAA 734

OY 167 GluThrProGlu-----LeuProLysProGlyValThrProArgSerGluVal 182

DB 733 ACAGCAGACGAGCCCGCTGCACATYTGACTCCGCCAGCGTCCCTCCGAC----- 683

OY 183 lIegluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSer 202

DB 682 -----GGGCGCGCAGCAGCAGCA 665

OY 203 AlalaSerSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu---Gly 221

DB 664 GCAGCTAC-----GAGCAGCAGCTACACACAGCAGCAGCTACAGCTACAGCAG 614

OY 222 LeuGluLysGlnAlaIleLysIleAspLysGluArgGluGlyGlnGlnLysAla 241

DB 613 CTACAGCAGCAGCAGCTACAGCAGCTACACACAGCAGCTACAGCAGCAGCAGCA 554

OY 242 AlaGluGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIle 261

DB 553 CTACACACACAGCTACACAGCAGCA-GCTACACACAGCAGCTACACAGCTACAGCA 495

OY 262 AlaValSerValAlaIleThrValIleSerIleValAlaAlaIleThrValIleThrValIle 281

DB 494 GCAGCTACAGCAGCTACACAGCAGCTACAGCTACAGCTACAGCTACAGCA 435

OY 282 GlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaAlaAlaAlaGlyGlyAlaAla 301

DB 434 GCAGCTACACACAGCAGCTACACACAGCAGCTACAGCAGCTACACAGCAGCAGCA 375

OY 302 GlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 321

DB 374 GCTACACACAGCAGCTACACAGCAGCTACACAGCAGCTACACAGCAGCTACACAGCA 315

OY 322 ValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAla 341

DB 314 GCTACAGCAGCAGCTACACAGCAGCTACAGCAGCTACAGCAGCTACAGCAGCT 255

OY 342 ValLysSerGlyIleLysAlaIleThrAlaIleLysValLysAlaIleLysAlaIle 361

DB 254 ACAGCTACACAGCAGCTACACAGCAGCTACACAGCAGCTACAGCAGCTACAGCAGC 195

OY 362 SerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPro 381

DB 194 AGGTCCTCGCTTCCACACACACCGGGTGGAGCAGTGAATGAAATATTTCAG 135

OY 382 LysLeuSerLysValIleSerSerLeuThrSerLysTrpVal 395

DB 134 GAAGCTTCATCCATCTTGATGAGGGGAGCAGCAGCTCTGGCTC 93

RESULT 4

CNS033D4

LOCUS

DEFINITION

CNS033D4 935 bp DNA linear GSS 15-MAY-2000

Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone 208D22 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION

AL225985.1 GI:7884892

VERSION

AL225985

KEYWORDS

GSS: genome survey sequence.

SOURCE

Tetraodon nigroviridis.

ORGANISM

Tetraodon nigroviridis

REFERENCE

1 (bases 1 to 935)

Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.

Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

TITLE

Unpublished

2 (bases 1 to 935)

Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.

Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL

Unpublished

3 (bases 1 to 935)

Roeest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.

Submitted (12-APR-2000)

Direct Submission

Genoscope.

COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

Location/Qualifiers

1. 935

/organism="Tetraodon nigroviridis"

/db\_xref="taxon:99883"

/clone="208D22"

/clone\_lib="G"

/note="Genoscope sequence ID : C0AG208DB11SP1-end : PUC-Or1"

BASE COUNT 136 a 251 c 273 g 271 t 4 others

ORIGIN

Alignment Scores:

Pred. No.: 6.45e-05 Length: 935

Score: 152.50 Matches: 62

Percent Similarity: 46.36% Conservative: 40

Best Local Similarity: 28.18% Mismatches: 105

Query Match: 6.44% Indels: 13

DB: 17 Gaps: 7

US-09-889-314-2 (1-496) x CNS033D4 (1-935)

OY 194 GlnThrLeuGlyGluAlaThrLysSerAlaLeuSerThrGlnAlaGln 213

DB 13 CAGGCCCTCGGTGTGCTGCTCAAAAGCTCCAAACATCTCGTCCGACCGCTCCAGG 72

OY 214 AlaAspGlnThrAsnLysLeuGluGlu---LysGlnAlaIleLysIleAspLysGlu 232

DB 73 AGAAACCAAGAAACCTTGTGAGCCAGAGCTGCTCCCTCATGCCAGATGATGAAGCT 132

OY 233 ArgGluGluThrGlnGlnLysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThr 252

DB 133 TCCTGAATATTCACACTACTGCTCCACCCCGTGGGTGTGGAACAGCAGCAG 192

OY 253 MetAspThrValAsnThrValAlaMetIleAlaValSerValAlaIleThrValIleSerIle 272

DB 193 CTGCTGAGCGCTACTGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 27, 2003, 14:10:57 ; Search time 1691 Seconds  
(without alignments)  
4750.421 Million cell updates/sec

Title: US-09-889-314-2  
Perfect score: 2369  
Sequence: 1 DTNMSISSSGPDNOKNIMS.....LKAVVAISGATAGAKHTNMF 496

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+P2n.model -DEV=xlh  
-O/cgn2.1/USP1000/US09889314/unaltd\_24012003\_144540\_10121/app.query.fasta.1.647  
-DB=EST -OPMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNIT=bits -STRAT=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdd -LIST=45  
-DOCCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=45 -MODE=LOCAL  
-OUTFMT=pro -NOM=ext -HEP5IZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09889314@cgn2.1.1.899@unaltd\_24012003\_144540\_10121 -NCPV=6 -ICPV=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-NARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em.estba:\*  
2: em.esthum:\*  
3: em.estin:\*  
4: em.estnu:\*  
5: em.estov:\*  
6: em.estpl:\*  
7: em.estro:\*  
8: em.htc:\*  
9: gb.est1:\*  
10: gb.est2:\*  
11: gb.htc:\*  
12: gb.est3:\*  
13: gb.est4:\*  
14: gb.est5:\*  
15: em.estfun:\*  
16: em.estom:\*  
17: gb.gss:\*  
18: em.gss.hum:\*  
19: em.gss.inv:\*  
20: em.gss.pln:\*  
21: em.gss.vit:\*  
22: em.gss.fun:\*  
23: em.gss.mam:\*  
24: em.gss.mus:\*  
25: em.gss.other:\*  
26: em.gss.pro:\*  
27: em.gss.tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161.5	6.8	936	17	CNS07CJY
2	155.5	6.6	1150	17	CNS079ZS
3	153.5	6.5	935	17	CNS033D4
4	152.5	6.4	935	17	CNS033D4
5	148	6.2	711	13	BM588321
6	147.5	6.2	619	17	BM588321
7	147.5	6.2	1094	17	CNS076CM
8	144.5	6.1	2295	17	BC016304
9	142	6.0	938	17	CNS07CG2
10	142	6.0	2524	11	BC028236
11	140	5.9	1845	11	AY103583
12	140	5.9	4272	11	BC016144
13	137.5	5.8	746	17	A2624776
14	136.5	5.8	987	11	AK013560
15	136	5.7	1194	9	AL334303
16	135.5	5.7	3344	11	BC028681
17	135	5.7	2589	11	BC028951
18	134.5	5.7	485	9	AA930444
19	134.5	5.7	809	17	CNS02310
20	134.5	5.6	1123	17	CNS06RUI
21	132.5	5.6	1027	11	AY107616
22	132	5.6	888	17	CNS075P4
23	131	5.5	892	17	CNS075P4
24	130.5	5.5	1101	17	CNS05HLZ
25	129	5.4	887	14	BO735450
26	128.5	5.4	1745	17	FR0007388
27	128.5	5.4	443	17	FR0008252
28	128.5	5.4	751	12	BG656282
29	128.5	5.4	1000	13	BM556446
30	128	5.4	619	17	FR0006944
31	128	5.4	711	13	BM588321
32	128	5.4	773	10	BE585195
33	127.5	5.4	494	10	AW660984
34	127.5	5.4	805	12	BG676639
35	127	5.4	644	14	BO838510
36	127	5.4	785	12	BG246146
37	126.5	5.3	1590	11	AK014874
38	126.5	5.3	1792	11	AY103948
39	126.5	5.3	2477	11	AK014336
40	126	5.3	786	13	BI106945
41	126	5.3	3110	11	AK019964
42	126	5.3	3441	11	BC030898
43	126	5.3	4088	11	BC021514
44	125	5.3	526	17	AZ312601
45	125	5.3	559	17	A2661219

ALIGNMENTS

RESULT 1	CNS07CJY	936 bp	DNA	linear	GSS 08-JUL-2001
LOCUS	T7 end of clone BD0A003E07 of library BD0A from strain CBS 94 of				
DEFINITION	Candida tropicalis, genomic survey sequence.				
ACCESSION	AL339268				
VERSION	AL339268.1				
KEYWORDS	GSS				
SOURCE	Candida tropicalis.				
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
REFERENCE	1 (bases 1 to 936)				
AUTHORS	Soulet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,				